Journal of
*Cybersecurity and Privacy*

MDPI

*Systematic Review*

# Humans and Automation: Augmenting Security Operation Centers

**Jack Tilbury and Stephen Flowerday ***

School of Cyber Studies, The University of Tulsa, Tulsa, OK 74104, USA; jack-tilbury@utulsa.edu
* Correspondence: stephen-flowerday@utulsa.edu

**Abstract:** The continuous integration of automated tools into security operation centers (SOCs) increases the volume of alerts for security analysts. This amplifies the risk of automation bias and complacency to the point that security analysts have reported missing, ignoring, and not acting upon critical alerts. Enhancing the SOC environment has predominantly been researched from a technical standpoint, failing to consider the socio-technical elements adequately. However, our research fills this gap and provides practical insights for optimizing processes in SOCs. The synergy between security analysts and automation can potentially augment threat detection and response capabilities, ensuring a more robust defense if effective human-automation collaboration is established. A scoping review of 599 articles from four databases led to a final selection of 49 articles. Thematic analysis resulted in 609 coding references generated across four main themes: SOC automation challenges, automation application areas, implications on analysts, and human factor sentiment. Our findings emphasize the extent to which automation can be implemented across the incident response lifecycle. The SOC Automation Matrix represents our primary contribution to achieving a mutually beneficial relationship between analyst and machine. This matrix describes the properties of four distinct human-automation combinations. This is of practical value to SOCs striving to optimize their processes, as our matrix mentions socio-technical system characteristics for automated tools.

**Keywords:** security operation center; security analyst; automation; automation bias; human–automation collaboration; levels of automation; automation characteristics

check for updates

**Citation:** Tilbury, J.; Flowerday, S. Humans and Automation: Augmenting Security Operation Centers. *J. Cybersecur. Priv.* **2024**, *4*, 388–409. https://doi.org/10.3390/jcp4030020

Academic Editors: Chrysostomos Stylios, Kyriakos Stefanidis and Vasiliki Liagkou

Received: 26 May 2024
Revised: 20 June 2024
Accepted: 28 June 2024
Published: 1 July 2024

## 1. Introduction

Despite the widespread adoption of automation in security operation centers (SOCs), security analysts remain overwhelmed by constant alerts. This problem is further compounded by the continuous integration of tools that generate additional security alerts with corresponding data and the existing shortage of skilled analysts to manage them effectively [1,2]. While the need for automation in SOCs is evident, its current implementation offers room for improvement. According to [3] (p. 03), 35% of 500 cybersecurity decision-makers surveyed admitted to missing a security alert, 22% stated that they had ignored security alerts entirely, and 25% confessed to not having acted upon a high-profile alert. These statistics underscore the potential consequences of automation bias and complacency amidst a top-heavy automation environment and elucidate the deficient quality of tool configuration and the lack of tool optimization. Automation complacency, the substandard automation monitoring, automation bias, and the tendency for individuals to become over-reliant on automated machines are significant concerns in SOCs. The earliest research on automation bias dates to the 1990s [4,5]. Despite automation bias and complacency not being novel concepts, limited efforts exist to examine them in the cybersecurity context [6–11]. This indicates the need for further attention, with this study taking initial steps to enhance the effectiveness of operations in SOCs by proposing human-automation collaboration strategies.

To achieve the desired benefits of SOC automation and keep the analyst-in-the-loop, optimal configuration, maintenance, and understanding of these tools must occur. The ineffective performance of automated tools often results from how they are programmed instead of supposed technological limitations [12]. Thus, automation is only capable of performing as instructed. Traditional automation relies on the approach where very specific rules are developed. These tools are typically employed to relieve the burden of manual inspection and detection. However, ref. [13] (p. 20217) assert that "automation with explicit rules has generally not been very successful because (1) there is an intrinsic difficulty in defining expert rules that robustly work in complex and dynamic networks, and (2) no single expert knows all the rules for different network technologies". This has given rise to artificial intelligence (AI) and machine learning (ML) solutions in security operations. Likewise, these tools can only perform based on how they are trained and how they learn. Therefore, automated tools and security analysts must co-exist with one another and strive to provide mutual benefit. This is supported by [14], who surveyed 641 analysts and found that partial automation (reliant on both the human factor and technological solutions) was the primary method of all threat-hunting activities. As SOCs increase the number of processes they automate, investigating how analysts interact with and gain advantage from using these tools represents an equally important avenue of research.

It is encouraging to see a strong call for further research in this domain [15]. When referring to analysts as the "human-as-a-security-sensor", ref. [16] assert that technology must amplify human capacity rather than diminish it. The authors call for studies investigating the degree of analyst inclusion during the design and implementation of automated solutions. The prevalence of automated cybersecurity tools should provide ample motivation toward studying human-automation interaction, yet research evaluating trust in these tools appears limited [9]. Finally, ref. [2] (p. 05) comment that "…the technical paths seem to be approached by researchers and developers, [and] the socio-technical aspects do not". Hence, integrating SOC automation that supports trust, limits bias, and does not encourage complacent behaviors constitutes a critically under-researched field.

This study aims to explore human-automation collaboration within the SOC environment. We have opted to use "collaboration" instead of "interaction" as we believe that modern forms of automation, particularly with the surge of large language models and decision support systems, extend beyond simple interactions. For instance, ref. [17] define human–automation interaction involving three elements: humans instruct automation on how to perform, humans can intercede with automation and control its actions, and humans can receive information from automation. Furthermore, in their study of automated agents in the incident response process, ref. [18] (p. 01) stated that "AI agents collaborate with humans as interdependent teammates to reach a common goal". Our discussion will emphasize the qualities that advanced systems must possess to elevate mere interaction into a collaborative relationship, sharing the sentiments of [19] (p. 185) who refer to human-automation collaboration as: "allowing for the creation of automation that can take 'tasking" or delegation instructions at a variety of levels and degrees of completeness while simultaneously providing feedback on the feasibility of those instructions … automation may not only report on the progress and outcome of delegated tasks but may also reason about what the human expects to occur during task performance-and thereby be in a better position to detect and report anomalies". Given this collaborative nature, the SOC can be described as a socio-technical system, constituting socio and technical parts that operate together in an environment and work towards completing tasks together in the most optimal manner [20].

To achieve our objective, we will provide a more granular focus on where (and how) automation is used in SOCs and consider its implications for security analysts. This study will also showcase the varying extent to which automation can be implemented. Furthermore, we will also consider the degrees of automation versus the levels of automation and provide a clear distinction between the two concepts.

## 2. Materials and Methods

This study followed the scoping review approach by [21]. The Preferred Reporting Items for Systematic Reviews extension for Scoping Reviews (PRISMA-ScR) checklist was adopted to inform the protocol for this study (see Figure 1).

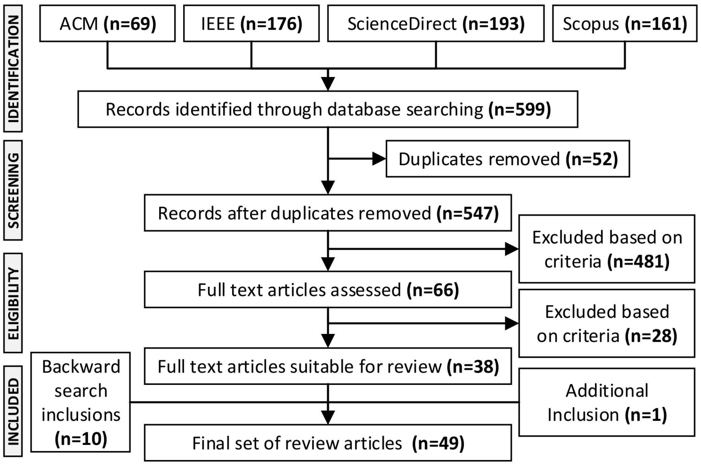
IDENTIFICATION
ACM (n=69)
IEEE (n=176)
ScienceDirect (n=193)
Scopus (n=161)
Records identified through database searching (n=599)
SCREENING
Duplicates removed (n=52)
Records after duplicates removed (n=547)
Excluded based on criteria (n=481)
ELIGIBILITY
Full text articles assessed (n=66)
Excluded based on criteria (n=28)
INCLUDED
Backward search inclusions (n=10)
Full text articles suitable for review (n=38)
Additional Inclusion (n=1)
Final set of review articles (n=49)

**Figure 1.** Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) diagram.

To search for relevant literature, we applied two search strings to the following four databases: ScienceDirect, ACM Digital Library, Scopus, and IEEE Xplore. The first search string concerned automation in SOCs, with the second factoring in the human component (for conciseness, Appendix A provides an example search string for a select database). This resulted in 599 papers, which was reduced to 547 after removing duplicates. The articles were then screened at the title and abstract level, where we followed the Population, Concept, and Context (PCC) framework put forward by [21]. Articles in this review needed to adhere to the inclusion criteria of a population of security analysts, the concept of automated security tools, and the context of SOCs. Based on this, 481 articles were excluded. The final 66 articles were read in full, and 28 articles were excluded on the grounds of not sufficiently addressing all three aspects. Backward searching (i.e., examining the reference list of included articles) resulted in 10 articles being included. Since running our database search, we recently published a 2023 conference article, indicating the one included article. Figure 1 displays our final PRISMA diagram.

## 3. Results and Findings

The study selection process yielded 49 articles published over the last 11 years. It is worth noting that no date limitations were applied when selecting literature, but it just so happened that the selected literature dates between 2013 and 2024. While the literature on SOCs before 2013 is available, discussions centered on how automation is used and its implications on analysts are part of an emerging field. Available data indicate that since 2019, there has been a rapid increase in literature examining the specificities of automation in security operations. Figure 2 shows the number of publications by year.

The final 49 articles were imported in NVivo, read in full, and thematically analyzed. The thematic analysis involves identifying themes and patterns in the literature and categorizing them accordingly [22]. We followed the six-phase approach, which produces both semantic (themes that directly present themselves in the literature) and latent (themes that are developed based on the interpretation of data) themes [22]. This resulted in 609 coding references being coded against 65 individual codes and further distributed across four main themes: (1) SOC Automation Application Areas, (2) SOC Automation Implications, (3) Human Factor Sentiment, and (4) SOC Challenges Necessitating Automation (seen in Figure 3).

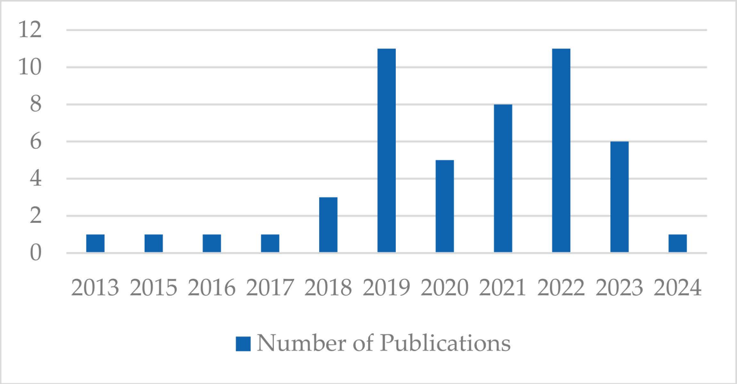

**Figure 2.** Number of publications by year.

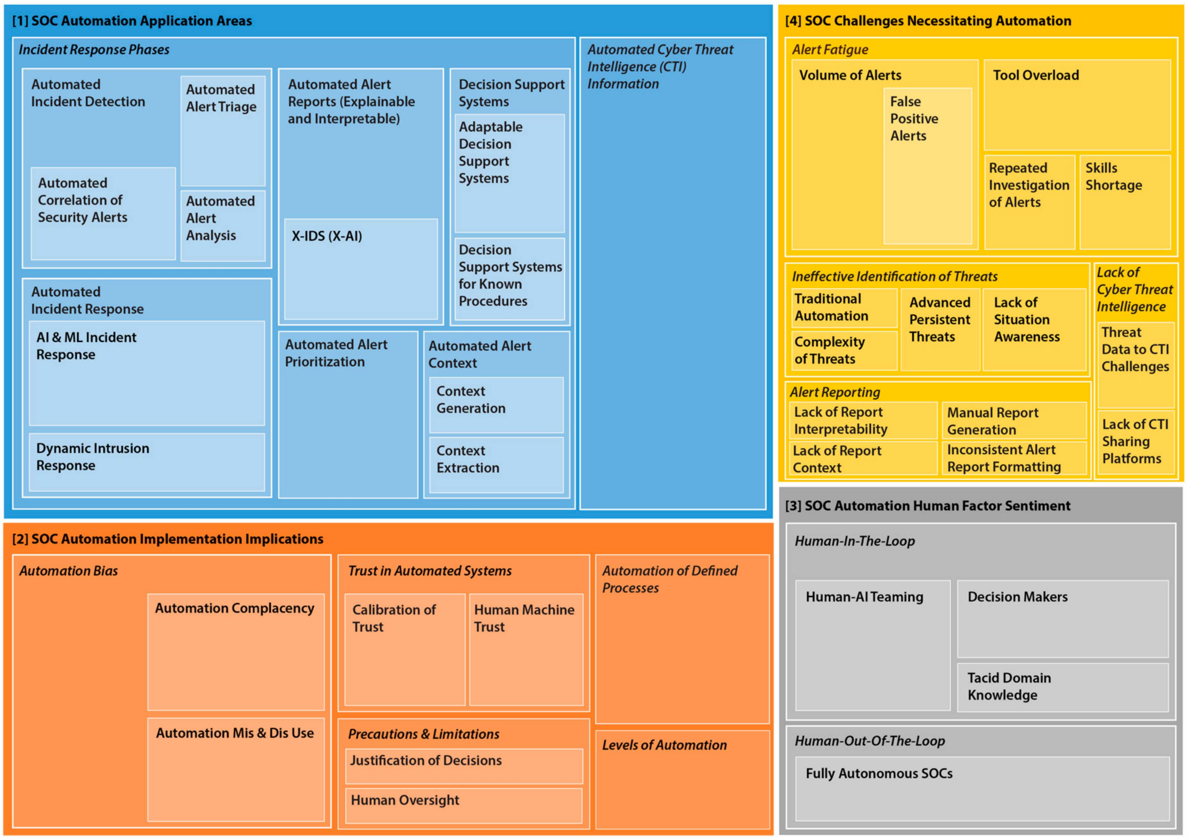
[1] SOC Automation Application Areas
Incident Response Phases
Automated Incident Detection
Automated Alert Triage
Automated Correlation of Security Alerts
Automated Alert Analysis
Automated Incident Response
AI & ML Incident Response
Dynamic Intrusion Response
Automated Alert Reports (Explainable and Interpretable)
X-IDS (X-AI)
Decision Support Systems
Adaptable Decision Support Systems
Decision Support Systems for Known Procedures
Automated Alert Prioritization
Automated Alert Context
Context Generation
Context Extraction
Automated Cyber Threat Intelligence (CTI) Information
[4] SOC Challenges Necessitating Automation
Alert Fatigue
Volume of Alerts
False Positive Alerts
Tool Overload
Repeated Investigation of Alerts
Skills Shortage
Ineffective Identification of Threats
Traditional Automation
Complexity of Threats
Advanced Persistent Threats
Lack of Situation Awareness
Lack of Cyber Threat Intelligence
Threat Data to CTI Challenges
Lack of CTI Sharing Platforms
Alert Reporting
Lack of Report Interpretability
Manual Report Generation
Lack of Report Context
Inconsistent Alert Report Formatting
[2] SOC Automation Implementation Implications
Automation Bias
Automation Complacency
Automation Mis & Dis Use
Trust in Automated Systems
Calibration of Trust
Human Machine Trust
Precautions & Limitations
Justification of Decisions
Human Oversight
Automation of Defined Processes
Levels of Automation
[3] SOC Automation Human Factor Sentiment
Human-In-The-Loop
Human-AI Teaming
Decision Makers
Tacid Domain Knowledge
Human-Out-Of-The-Loop
Fully Autonomous SOCs

**Figure 3.** Thematic analysis coding strategy result.

Figure 3 displays the result of the final coding process, with each block illustrating the proportion of each theme in relation to the 609 coding references. For instance, automation application areas mark the most dominant theme discussed in the literature. Similarly, there is room for more research to address the human factor sentiment in SOCs, discussing analysts' tacit domain knowledge, their role as decision-makers, and how to best team them with AI tools and technologies.

## 3.1. SOC Challenges Necessitating Automation

Given that prior works have identified these challenges, [15,23–25], coupled with the case for article brevity, we will not delve into an exhaustive explanation of each challenge. Through our systematic approach, we categorized SOC challenges into four distinct themes:

(1) lack of cyber threat intelligence, (2) ineffective identification of threats, (3) alert fatigue, and (4) alert reporting constraints. The thematic map is shown in Appendix B.

### 3.1.1. Lack of Cyber Threat Intelligence

A core issue was the lack of adequate solutions to convert threat data into threat intelligence. Ref. [26] (p. 02) define threat intelligence as "evidence-based knowledge, information or data about existing or emerging threats that can be used to prevent or mitigate them". Additionally, ref. [27] (p. 149) define threat intelligence as "evidence-based knowledge, including context, mechanisms, indicators, implications and actionable advice, about an existing or emerging menace or hazard to assets that can be used to inform decisions...". Simply collecting threat information does not assist security analysts in learning about and responding to threats [28]. In the form of actionable intelligence, alert meaning must be derived, the data redundancy removed, and the information must be structured. Given the amount of threat information collected, identifying novel threats amongst this data can be a cumbersome endeavor [29]. Ref. [30] contend that where automation can assist in scripting rules for intrusion detection systems, security analysts are further responsible for identifying the threats to which the rules pertain. Once threat intelligence had been extracted, studies expressed their dissatisfaction with the lack of integrated and automated threat intelligence information-sharing platforms across the industry [31,32]. These challenges were mentioned in 8 out of 49 (16%) articles analyzed.

### 3.1.2. Ineffective Identification of Threats

The inability of traditional forms of automation (intrusion prevention systems and intrusion detection systems) to detect novel threat variations was cited as the predominant reason for cyberattacks materializing [33,34]. This appears to be a twofold issue. Traditional automation is largely signature and anomaly-based, and threats are becoming more complex in their execution. Ref. [35] issue remarks that complexity (as seen in advanced persistent threats) results in increased indicators of compromise that security analysts and their tools must now account for, leading to an additional administrative burden. Attributing ineffective detection to poor rule configuration and tools repeatedly issuing unsolicited alerts was also reported [36]. Considering that IDSs were initially proposed in the 1980s, ref. [37] encourage moving beyond passive cybersecurity (one that is static) and into the realm of active defense through the implementation of AI and ML technology. Anomaly-based detection systems generate security alerts because the flagged behavior differs from what is usually observed. However, refs. [38,39] mention that abnormal behavior is often not related to maliciousness. Cyber situational awareness has been defined as gathering information about a situation and utilizing it to differentiate between the suitability of potential actions [40]. A security analyst's failure to establish situation awareness in a SOC due to the multitude of responsibilities they have (including monitoring, detecting, responding, configuring tools, integrating new tools, team, and AI-agent collaboration) constitutes the failure of effective threat detection according to two papers [2,41]. These challenges were mentioned in 18/49 (37%) articles analyzed.

### 3.1.3. Alert Fatigue

The most well-known challenge that plagues SOCs is the volume of alerts that their analysts must triage [24,42,43]. Reasons for this include an increased number of tools, poor configuration, and anomaly detection systems, among others. This issue is exemplified by [39] who reported a true positive alert ratio of 1.59% when analyzing an alert dataset. A frequent discussion point is that of false positives, which [44] define as genuinely flagged security alerts that do not have an associated security event and are usually caused by loosely defined rules. According to the same authors, false positives are differentiated from benign triggers, which are true alerts that match a configured rule but are purposefully ignored (due to legacy systems, for example). The voluminous alert issue has knock-on effects. Firstly, not only does the flood of alerts need to be analyzed, but this information must also be stored and backed up. Secondly, the timely nature of response forces analysts to make quick judgments concerning which alerts require a deeper investigation [43]. Furthermore, human factor implications of desensitization after repeated exposure and a reduction in response accuracy as a result of an overwhelming amount of information processing are mentioned [39,45]. Correlating with industry reports [3], security analysts state that the constant integration and maintenance of security monitoring tools has significant secondary effects, namely, how time-consuming it is and the resultant increase in alerts [31,46]. Refs. [36,47] both address how the continuous addition of tools can lead to analysts becoming overwhelmed and suffering from alert fatigue. Another study commented on the tedious tool integration process, given the number of APIs (averaging over 200 APIs per tool) and plugins that need to be accommodated [48]. Investigating the same intrusion, either in parallel by different analysts or sequentially by the same analyst, results in individual alerts being assessed in isolation instead of analyzed collectively concerning the same event. However, ref. [36] claim that not all tools can group associated alerts, rendering a large volume of alerts and their repeated investigation inevitable. The repeated investigation of alerts was also mentioned by [1] who found that security analysts often inspect the same domain on different occasions without arriving at a final solution. These challenges are amplified by the skills shortage that security operation centers are currently subject to [11,49]. These challenges were mentioned in 31/49 (63%) of the articles analyzed, making alert fatigue and its sub-themes the most dominant issues experienced by security analysts.

#### 3.1.4. Alert Reporting Constraints

Studies suggested that alert reports lacked context and were inconsistently formatted. One study indicated that security analysts face the cognitive task of piecing together the fragments of intrusion information while still considering the broader organizational impact [35]. Further literature reports that a lack of contextual information elongates response initiatives as analysts need to source information not presented to help inform their response decisions [49,50]. Thus, the process of gaining contextual knowledge and generating actionable information to accompany current alert reports is largely a manual undertaking [51]. Additionally, the lack of industry and tool standardization regarding alert reports makes collecting and interpreting information more difficult [31]. The technical jargon of alert reports inhibits analysts from a comprehensive understanding, with studies characterizing alert reports as lacking human-readable information. Conversely, [32] argue that report information should be machine-readable to automate responses. These challenges were mentioned in 12/49 (24%) articles analyzed.

### *3.2. SOC Automation Application Areas*

We must acknowledge where automation is employed in SOCs (its application areas) to address security analysts' challenges and achieve a collaborative human–automation environment. This theme considers the current use cases of SOC automation and mentions novel solutions that are being proposed (see the thematic map in Appendix C).

#### 3.2.1. Automated Cyber Threat Intelligence (CTI) Information

With the rise of threat information-sharing platforms, vulnerability databases, and internet-based cybersecurity information, SOCs are challenged with converting large volumes of threat information into threat intelligence. Considering big data, automated solutions are needed to collect and classify information accordingly. Such processes can aid in quickly educating SOC analysts. A commonly mentioned use case in this area is automated rule generation for intrusion detection systems (IDS), whereby tools automatically collect CTI, generate indicators of compromise (IOCs), and update IDS for the detection of malicious activity [26,29–31,52]. In one study, more than 70% of analysts stated that automated tools must improve the quality of the information they collect [31]. Presenting security analysts with comprehensive information that can be effectively leveraged represents a core goal of these tools. Another study recommended collecting threat intelligence from malware information-sharing platforms (MISP) instead of general web data to achieve data quality and developed SmartValidator, a tool that continuously and automatically collects CTI and validates internal security alerts against it (i.e., proving their legitimacy) [30]. They showcased how SmartValidator performed better when MISPs were the primary data source. Ref. [52] posited that the manual generation of IOCs is more interpretable and human-friendly, allowing SOC analysts to understand them sufficiently. Thus, they compared their automated IOCs to manually generated ones and noted that security analysts found no difference between the two sources. The notion of threat intelligence interpretability was also discussed by [29] who stated that security analysts must efficiently identify and comprehend all the relevant information when analyzing automated threat reports.

**Findings:** Overall, we classify automated CTI tools operating in four sequential phases:

1. Collect threat information.
2. Analyze and convert threat information into threat intelligence.
3. Validate the cyber threat intelligence against internal alert information to determine the presence or absence of intrusions.
4. Generate indicators of compromise (IOCs).

By scanning the external threat landscape for information that can assist SOCs in efficiently detecting threats, automated CTI tools permit security analysts to channel their efforts into effective decision-making strategies [30]. These four phases illustrate that the automation of CTI information can be implemented to varying degrees—with the recommendation of all four phases being adhered to.

### 3.2.2. Automated Incident Detection

Where applicable, we attempted to code SOC automation application areas by the incident response lifecycle (IR) phases. We focused on opposite spectrums of the IR lifecycle, namely detection and analysis at the one end and recovery and response at the other. This was performed because traditional cybersecurity automation is typically employed in the earlier phases. However, as technology progresses, particularly with the rapid evolution of AI and ML tools, new recovery and response solutions are beginning to emerge. For example, Microsoft's Security Co-Pilot and CrowdStrike's Charlotte AI (neither mentioned in any reviewed articles) represent first-to-market large-language models for use in SOCs.

One of the essential components of harnessing automation within SOCs is to support analysts in identifying threats by providing them with the highest priority security alerts to act upon [53,54]. Automation that specifically flags the most severe threats (those demanding immediate action) leads to beneficial efficiency gains. For this reason, we coded the sub-theme of automated "alert analysis." Here, our results yielded two categories of solutions: the automation of the alert correlation process [49,55,56] and the automation of the triage process [43,47,48]. One study developed a semi-autonomous tool interacting with security analysts to identify advanced persistent threats [57]. Another study aimed to detect malware attacks by monitoring host machines and their potential communication connections with malicious domains hosted at command-and-control centers [54]. Through 24 interviews conducted with SOC analysts, the effectiveness of automated tools identifying suspicious activity primarily depends on their configuration [12]. Based on a mixed-methods approach with 20 SOC analysts, ref. [44] designed a framework for configuring tools to generate reliable, explainable, analytical, contextual, and transferrable alerts. Various authors suggested corroborating alerts in conjunction with the preceding and proceeding alerts (i.e., the alerts around the flagged alert) [41,51,56,57].

**Findings:** Overall, in many of the examined articles, automation was commonly used for three primary purposes in the detection phase:

1. Connecting and correlating alerts from different security tools integrated into the SOC.

2. Gathering suspicious alerts in context with their surrounding (and related) alerts to provide richer context.
3. Grouping alerts based on similar characteristics and IOCs (often called clustering). For instance, ref. [55] advised that automation should analyze flagged alerts in context with similar previous incidents to allow analysts to determine how to remediate them.

These three purposes illustrate that the automation of incident detection can be implemented to varying degrees. It is recommended that all three aspects be adhered to.

#### 3.2.3. Automated Alert Context

Once threats and vulnerabilities have been detected, the next step is to supplement these alerts with additional context. Truly aiding security analysts extends beyond simply flagging alerts as malicious. It involves providing further background information about these alerts in the form of alert characteristics and properties, where they have originated from, attacker behaviors and techniques, and what the alert impact could be if not resolved [45]. Ref. [58] implemented Context2Vector to reduce security analysts' burden during the triage process by enhancing alerts' informative information beyond what traditional intrusion detection systems offer. In the cybersecurity domain, ref. [45] (p. 02) define alert context as "a comprehensive concept, cover[ing] different kinds of data sources, such as the provenance . . ., the logs collected within the same period, [and] the alert sequences".

**Findings:** Within this theme, our analysis led us to interpret that alert context is first extracted and then generated:

1. Context extraction refers to collating organizational information affected by security alerts (i.e., pulling context from additional sources beyond the SOC). The more information that is collected from around the organization, such as business assets affected and processes disrupted, the better decision-making will be [58,59]
2. Context generation refers to utilizing and generating the auxiliary extracted information and generating from it, such as the threat's potential impact and risk level. As a result, prioritization techniques are ameliorated [44,45,50].

The multi-phase approach of context extraction and generation illustrates that the automation of alert context can be implemented to varying degrees, with the recommendation that both phases be adhered to.

#### 3.2.4. Automated Alert Reports (Explainable and Interpretable)

The previous three themes have shown that, until now, threat information from the external environment has been collected, indicators of compromise have been created, intrusion detection systems have been updated, alerts have been detected, and context has been extracted and generated. As mentioned earlier, the next logical step in the incident response lifecycle is the culmination of the information into alert reports presented to analysts. A significant finding in the articles evaluated is that alert reports must build upon the contextual information gathered by creating a storyline of intrusion events [12,18,37,38,47,51]. By knitting this information together, security analysts are presented with different aspects of events in a timeline format, dating back to the root cause [54,56,57]. The premise of these reports is to explain what has happened, why it occurred, and its importance when analyzing a suspicious alert [38]. Therefore, for security analysts to derive value from alert reports, they must be able to comprehend and action them. Furthermore, an understanding of how these reports are developed is needed for security analysts to trust the systems generating the alert reports. [37] (p. 112394) refer to this as X-IDS (or explainable intrusion detection systems) and X-AI (in the context of AI), defining it as "the system's ability to explain the behavior of AI-controlled entities".

**Findings:** Drawing upon the discussion by [37], we have produced the diagram in Figure 4 to help distinguish between the terms interpretability, explainability, and understandability in the context of automated systems.

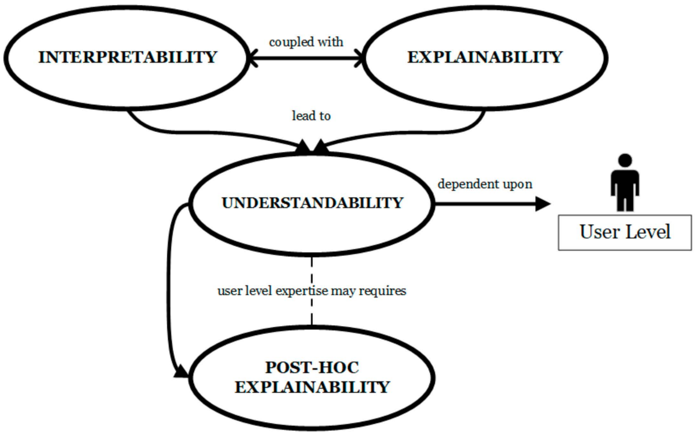

**Figure 4.** X-AI terminology.

1. Interpretability refers to automation's ability to produce information that security analysts can comprehend and understand. In SOC automation, interpretability occurs at the alert level, referring to the security event flagged. It can be measured by the analyst's ability to piece the intrusion storyline together.
2. Explainability occurs at the system level, whereby automated tools provide analysts with an explanation and justification of how they operate and why they have arrived at a particular decision. Ref. [37] (p. 112394) define explainability as "following the reasoning behind all the possible outcomes [whereby] models shed light on the model's decision-making process." Hence, explainable automation makes significant efforts to keep the human-in-the-loop. While these two constructs can exist independently, we recommend that they be coupled together, as seen at the top of Figure 4.
3. Sufficient interpretability and explanation lead to increased understanding. However, the understandability of this information depends on the audience and their skill level. Ref. [51] state that information presented to different stakeholders must adhere to their expertise and terminology.
4. If the user-level expertise cannot comprehend the information already presented, post hoc explainability in the form of data visualizations can be presented [37] but may come with performance issue trade-offs.

The distinction between these terms is also supported by [44]. The four-construct model in Figure 4 illustrates that the automation of explainable and interpretable alert reports can be implemented to varying degrees, with the recommendation that all four concepts are adhered to.

### 3.2.5. Automated Incident Response

Various automated solutions were present in the incident response phase, demonstrating automation's technical advancements. Examples include decision support systems for known threats (rule-based systems) [35,59] and decision support systems for unknown threats (adaptable intrusion response systems) [60,61]. Concerning decision support systems, ref. [62] advocate for semi-autonomous systems that either offer recommendations and implementations depending on the threat faced. Several authors promote dynamic intrusion response initiatives—ones that are not solely based on predefined configured responses from security analysts but cater to variations of attacks [56,60]. For instance, ref. [32] developed a dynamic intrusion response database that can be queried for the most recent countermeasures concerning a specific vulnerability. Ref. [26] differentiate between indirect automation integration and direct automation integration. The former assists in informing security analysts about threats and how to respond to them, and the latter will sense cyber threats and automatically respond to threats without human intervention. The decision to recommend and automatically implement a response is driven by the contextual information collected [58]. This includes associated business processes and assets, instantiating the importance of context extraction and generation.

Furthermore, the advent of AI and ML solutions in SOCs is clear [18,28,36,63]. 10 SOAR platforms were reviewed in one study, and it was and revealed AI/ML's use cases across the incident response lifecycle phases [55]. For instance, IBM's Resilient QRadar SOAR integrates ML models trained using historical data to predict the severity of new incidents, estimate the time for resolution, and link similar recently resolved incidents. The rise of large language models has also extended into SOCs, e.g., Microsoft's Security Co-Pilot and CrowdStrike's Charlotte AI. Regarding Microsoft's tool, it is worth noting that in and of itself, the term "co-pilot" is synonymous with assisting or helping, showcasing the tool's effort to augment SOC analysts' capabilities.

However, given that generative AI and ML tools require users (i.e., SOC analysts) to feed information into them, the confidentiality, integrity, and availability of an organization's security intelligence information (and to whom it is exposed) may hinder adoption rates. Ref. [18] evaluated security analysts' comfort levels with AI agents during the response phase and found that it was significantly lower when compared to the identification phase (M = 2.71 and M = 4.04, respectively, $p < 0.001$). More than half of the studies (29/49 or 59%) included in this analysis mentioned AI and ML solutions. Within the AI and ML domain, we found that 24.13% of articles focused on neural networks, 13.79% on visualization modules, and 10.3% on natural language processing (see Table 1).

**Table 1.** Articles by Technology Type.

| Technology Type | Number of Articles |
|---|---|
| AI/ML | 29 |
| Not Classified (simple AI/ML) | 12 |
| Neural Networks | 7 |
| Visualization Modules | 4 |
| Natural Language Processing | 3 |
| Deep Learning | 2 |
| Reinforcement Learning | 1 |
| IDS | 3 |
| SIEM/SOAR | 3 |
| Other | 5 |
| N/A | 9 |

**Findings:** AI and ML solutions in the incident response phase must be accompanied by contextual information and data visualizations to aid analysts' understanding of how these systems arrive at solutions [58]. Furthermore, when these systems are utilized, it is imperative to know when they are better off as recommendation engines or implementation agents. Also relevant to the discussion of AI and ML models is whether they are subject to supervised or unsupervised learning, a topic that sparked much debate within the literature reviewed [36–38,64]. We conclude that the critical nature of missed alerts demands highly trained supervised learning models to ensure greater accuracy in detecting threats. Nevertheless, due to the unpredictable nature of cybersecurity and its complexity, security analysts can gain advantages from using unsupervised models that detect abnormalities and patterns beyond traditional tool recognition. Hence, we suggest adopting a semi-supervised approach. The varying levels of autonomy illustrate that employing AI and ML solutions in incident response can be completed to varying degrees, with the recommendation that this be applied on a case-by-case basis. To better understand when to harness automation in the response phase and the level of automation that should be used, we recommend applying the four critical success factors put forward by [65]: (1) task-based automation, (2) process-based automation, (3) automation performance appraisal, and (4) SOC analyst training of automation systems.

In our findings, this discussion has highlighted automation areas of prominence and delineated their associated system characteristics. This information has been summarized in Figure 5.

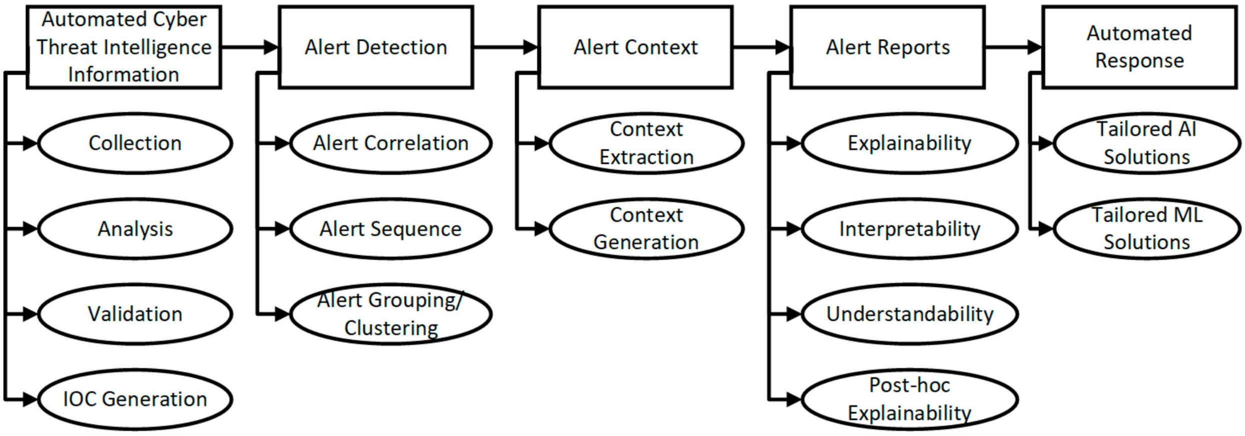

**Figure 5.** Automation application areas and system characteristics.

### *3.3. SOC Automation Implications on Analysts*

Limited studies have empirically measured automation bias and complacency within SOCs while discussing the ideal level of automation that should be applied. Consequently, we did not produce an individual thematic map for this theme. Instead, we report on results per the bottom left-hand rectangle of Figure 3. The same approach was taken for the human factor sentiment theme, referencing the bottom-right rectangle of Figure 3, respectively. The influx of SOC automation increases analysts' susceptibility to automation bias and complacency. Several discussions focused on automation disuse (the purposeful neglect and under-utilization of automation) and misuse (the over-reliance on automation). Refs. [6,7] found the occurrence of misuse whereby participants placed high trust in a poorly performing system. Another study suggests that automation errors can result in disuse because analysts initially perceive these tools to be error-free [53]. However, disuse can lead to legitimate attacks going undetected or ignored [34]. In combating automation bias, studies comment that accuracy cannot be the only metric evaluated during tool development and integration. To instill analyst confidence, automation must disclose confidence levels in its decisions [29,36]. Another study extended this by promoting an interactive environment where analysts can query why specific actions have been taken [28]. Proposing the Influence, Impact, and Mitigation of the Automation Bias model [65] contends that security analysts fall victim to automation bias because of the path of least cognitive effort, the diffusion of responsibility, and the perception of superiority and authority. Furthermore, they argue that automation bias can lead to a loss of situational awareness, cognitive skill degradation, task complacency, and a lack of vigilant information seeking.

**Findings:** To avoid bias and complacency, trust in automation must be calibrated so that well-performing tools are utilized accordingly and inconsistent tools are treated apprehensively. Factoring in the findings from previous sections showcases that providing clear interpretations of events to analysts will assist in preventing situations in which objectively well-performing tools are underutilized (i.e., preventing automation disuse). Additionally, the exposure to bias and complacency can be reduced by applying higher levels of automation to well-defined processes and against well-modeled, less severe threats. Contrary to that, ambiguous processes must be overseen and managed by security analysts [12]. The automatic execution of mitigation and restoration strategies is increasingly recognized in the literature. For instance, ref. [13] introduced an Action Recommendation Engine (ARE) that continually monitors network systems, offering tailored recommendations for actions to respond to identified malicious traffic. Based on raw data, alert reports, system feedback from previously recommended actions, and the severity of the threat, the engine is built to either provide an action recommendation or directly apply the action itself. When following the latter approach, the authors acknowledge that this leads to the "human-out-of-the-loop". The action recommendation engine illustrates the degree to which levels of automation can be applied.

### *3.4. SOC Automation Human Factor Sentiment*

Security analysts are important because of their decision-making skills and tacit domain knowledge. One way to partially mitigate automation is to limit where automation is applied. One study participant mentioned that "automation recommendations are fine, but the final decision should be on the incident handler," advocating for limited (if any) automation in the response phase [61] (p. 16). Another study credited security analysts for their ability to handle intuitive tasks, with automation better catered to resource-intensive repetitive tasks [62]. Studies suggested that the SOC marks a prime environment for human-AI collaboration whereby analysts can train automation and automation can improve efficiency for analysts [37,39].

**Findings:** To achieve a mutually beneficial relationship between security analysts and automation, involving the analysts during the development and integration of new tools is critical. Automation will only be as skilled as the analysts who program it. Therefore, recognizing the domain knowledge that analysts possess, specifically concerning Tier 3 analysts, will result in more intelligent technical solutions being identified. It is important to note that an organization will derive little benefit from an environment purely analyst or automation-based. Instead, the two entities must complement each other.

## 4. Discussion

Throughout the results and findings, it was repeatedly shown that automation can be implemented to varying degrees. In addition to that, our findings in the SOC automation implications on analysts' section revealed that varying levels of automation can be applied (i.e., only to provide recommendations or to action-specific responses autonomously). It is worth distinguishing the difference between these two concepts: degrees of automation and levels of automation:

1. Degree of Automation: The number of processes that a SOC automates. For example, SOCs with high degrees of automation will have automated processes throughout the incident response lifecycle. SOCs with low degrees of automation may only automate the detection of alerts. This can also be described as the breadth of automation.
2. Level of Automation: The level of autonomy that automated SOC processes possess. For example, high levels of automation in the response phase will recommend and implement response actions, with the analyst's option to intercede if necessary. Conversely, lower levels of automation in the response process may only provide analysts with several alternative response strategies. This can also be described as the depth of automation.

Therefore, a high degree of automation does not necessarily equate to high levels of automation. This is supported by [66] who define automation's breadth as the number of tasks and contexts that automation can handle and define automation's depth as the level of improvement that automation can achieve on specific tasks.

### *4.1. The SOC Automation Matrix*

The SOC Automation Matrix (see Figure 6) represents our main contribution. Our solution drew inspiration from the structure of the Boston Matrix, but it is important to note that it reads differently. Where the Boston Matrix analyzes a portfolio of products worthy of continued investment [67], the SOC Automation Matrix requires continued investment until optimization is reached. This is because SOCs are essential to modern-day organizations. Therefore, it is a matter of how efficient and effective a SOC can be rather than whether an organization should have one. Reading from left to right (x-axis), the automation present within a SOC increases. Our matrix will address both the degrees and levels of automation implemented in SOCs. Reading from bottom to top (y-axis), the degree of SOC analyst involvement increases. Each of the four quadrants is marked by the filled orange circle in the top left corner and discussed in terms of effectiveness and efficiency: Effectiveness concerns the human factor skillset and efficiency concerns automation.

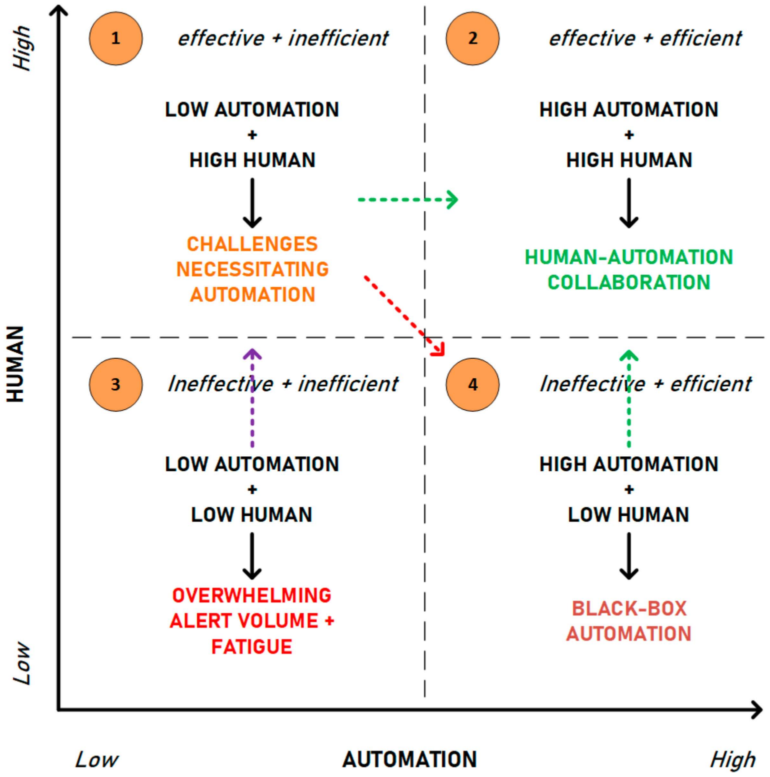

**Figure 6.** SOC Automation Matrix.

### Quadrant 1—Low Automation + High Human

This quadrant represents where most SOCs currently are due to manual, timely, and repetitive processes that are over-reliant on human analysts. Here, it is stated that activities are effective in that skilled SOC analysts tackle issues and apply their contextual knowledge to threat mitigation. However, their effectiveness alone is only valid until a certain point, whereby complex alerts and sophisticated attacks require additional competencies. Additionally, efficiency is lacking as the volume of alerts exceeds the resources available to cope with them. As SOCs grow in the information they monitor, alert fatigue presents itself as the central challenge, with others such as ineffective identification of threats and lack of cyber threat intelligence information also occurring, necessitating automation. SOCs in this quadrant do not have sufficient automation implemented. Reasons for this could vary, including a lack of trust, unwanted technological reliance, or financial limitations. This quadrant is characterized by automation disuse, i.e., the under-utilization of automation where it would otherwise be beneficial. The current narrative within a large body of literature states that the solution to all SOC challenges is full automation (i.e., integrating more automation into the SOC). Based on this, ill-advised SOCs deem their next logical step to move to Q4 (high automation + low human)—marked by the dashed red arrow in Figure 6. However, we suggest that instead of jumping diagonally down to Q4, SOCs strive for lateral movement toward Q2 (high automation + high human)—marked by the dashed green arrow exiting Q1.

### Quadrant 2—High Automation + High Human

This quadrant represents the most desirable state that SOCs should strive towards, as sufficient automation exists to assist SOC analysts. Here, activities are effective in that skilled SOC analysts are teaming with intelligent automated agents in threat mitigation and efficient in that automation takes care of routine and manual operations. Consequently, the human analyst is efficient and effective, and so is automation. Overall, tasks are performed swiftly and more accurately. The benefits of socio-technical system design are realized in Q2, and the SOC design is intended to "optimize the organization's social and technical systems jointly" [20] (p. 459). Strong human-automation collaboration exists in this quadrant, leveraging each entity's skills. An applicable example of such a tool is DeepLog, proposed by [68], which organizes and structures vast amounts of log data. This is achieved by automatically collecting logs generated concurrently, ordering them sequentially, grouping them according to their workflow, and separating them based on the task or event they relate to. Similar log mining and analysis techniques are mentioned in [69]. This allows analysts to identify the current point of threat execution better and conduct root cause analysis. Furthermore, the tool extracts contextual information beyond commonly collected data (log keys, for instance) to give analysts a complete understanding of the event. DeepLog represents an example of a mutually beneficial human–automation relationship in that it incorporates analyst feedback in situations where log entries were incorrectly classified as anomalies [68]. The premise of this quadrant is that neither entity overpowers the other. The system characteristics of automation (as per Figure 5) are fully implemented (see Figure 7). However, the potential for automation bias exists depending on the degree of automation implemented and the level of automation applied.

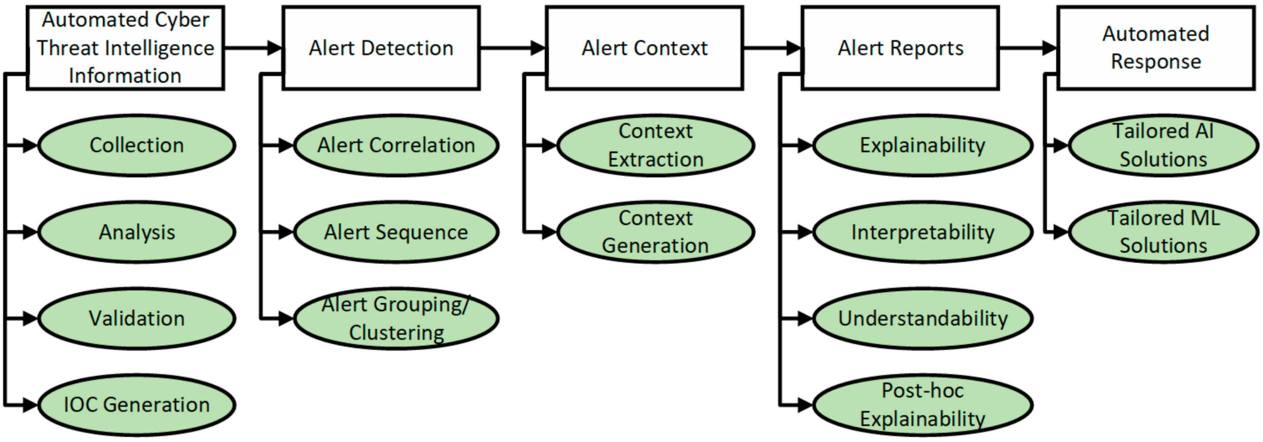

**Figure 7.** Fully implemented system characteristics.

### Quadrant 3—Low Automation + Low Human

This quadrant represents the least desirable state for SOCs as there is insufficient human involvement and automation implementation. Here, activities are ineffective because there is a lack of skilled SOC analysts and inefficient, as the lack of human resources is coupled with the lack of assistive automation. The few analysts that are present quickly become overwhelmed by the volume of alerts. Privacy fatigue, defined by [70] as the process whereby users experience exhaustion and cynicism about preserving privacy over an extended period is likely to occur. In the instance of SOC analysts, this fatigue represents itself as alert fatigue. Improper management of alerts ensues, and there is an increased potential to miss threats. A move to any other quadrant would be better for SOCs in this position but would require substantial investment. Therefore, to better understand what tools are required to match the needs of the SOC in question, skilled analysts should be advised. Moving vertically up to Q1 is recommended—marked by the dashed purple arrow in Figure 6.

### Quadrant 4—High Automation + Low Human

This quadrant represents a step in the right direction for SOCs but lacks complete optimization due to the primarily discredited human factor. Here, activities are partially ineffective in that automation's abilities are prioritized ahead of skilled SOC analysts. Efficiency gains are realized due to the degree of automation implemented. A relevant example of such is discussed by [69], who introduces SCOUT to identify anomalous outliers by grouping streams of log data into their relevant activities. While there exist encouraging system features such as grouped log examination (on the premise that alerts/logs may appear benign when analyzed individually, but when examined in conjunction with the events in which they were related, malicious intent can be discovered), the authors state that SCOUT operates on unsupervised learning models and can operate without human interaction. Thus, we posit that this tool is better positioned in Q4 than in Q2. It was mentioned that to appropriately classify activities and identify threats, domain experts (i.e., skilled security analysts) must be able to differentiate between well-modeled malicious and benign actions and rely on automated systems to identify threats and classify activities. Thus, reliance on automation is evident and can be justified if the automation is explainable, providing reasoning to analysts [37,69].

Automation in Q4 will assist in CTI information, incident detection, alert context, reports, and response. However, the system characteristics of automation (as per Figure 5) are limited. Only the high-level system characteristics are implemented, marked by the green-shaded characteristics in Figure 8. Moreover, despite the presence of automation, it is incorrectly leveraged, and it is likely to possess restrictive characteristics (see Appendix D) and be dominated by the identified challenge of alert reporting constraints. These lead to behavioral changes whereby the SOC analysts do not learn from automation but adapt their behavior to do what automation says. At this point, there is too much trust in automation and its ability, increasing the potential for automation complacency. As with Q2, there exists a possibility of automation bias. Trust must be calibrated by implementing the remainder of the system characteristics. Hence, SOCs in Q4 are encouraged to move vertically up to Q2—marked by the dashed green arrow exiting Q4 in Figure 6.

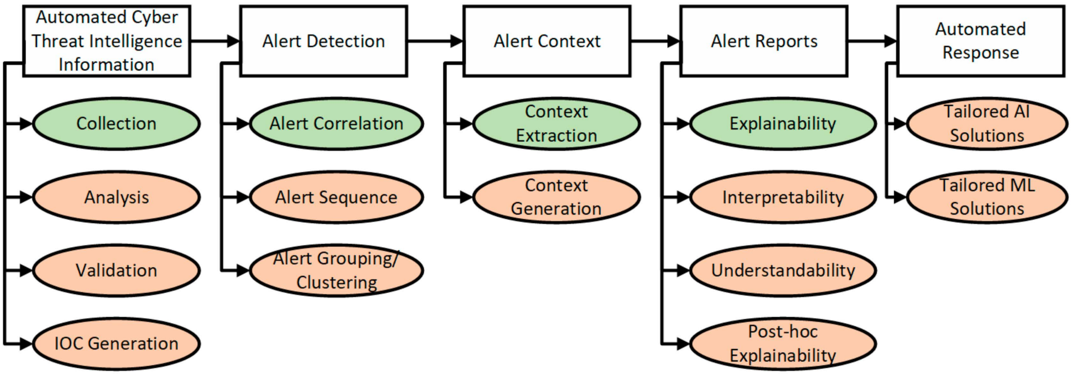

**Figure 8.** Partially implemented system characteristics.

To derive practical application from the matrix, SOCs can view the four quadrants as a maturity model that outlines their path toward efficient and effective practices. By mapping their challenges against those mentioned in the paper, real-world SOCs can identify that they likely fall on the left-hand side of the matrix. Conversely, suppose SOCs are confident that they have a high degree of automation implemented (regardless of its characteristics). In that case, they can identify that they likely fall on the right-hand side of the matrix. SOCs in this position should ultimately strive to be in Q2 over Q4 and can achieve this by working toward implementing the 15 system characteristics shown in Figure 7. In summary, SOCs begin in Q3 and are classified as being in an "initial and immature" phase. As SOCs grow and become better managed, they move into Q1. Their challenges cause them to rush toward implementing increased automation, finding themselves in Q4. Here, they are more defined in their operations, but SOC analysts' learning is prohibited. As SOCs mature, realizing they need to take advantage of analyst and automation skill sets, they work toward Q2, whereby optimization occurs.

### *4.2. The SOC Automation Matrix—Considering Levels of Automation*

Making use of the level of automation (LOA) framework put forward by [71] (comprising ten levels of human–automation collaboration delineating the split of responsibilities between human operators and machines), levels 2–9 split the responsibilities of humans and automation, respectively, with levels 1 and 10 being the complete responsibility of either entity. A SOC may possess a high degree of automation (Q2 and Q4) but exercise more control in terms of its human responsibility and opt for level 3—where automation throughout the IR lifecycle assists in determining available options and providing suggestions that analysts are not required to follow [72]. Alternatively, the same SOC may opt to automate many processes at level 8—whereby automated tools perform the entire job and inform the analyst what was performed if prompted.

We have illustrated this in Figure 9, focusing on Q2 and the levels of automation that could be applied. We can visually portray this phenomenon by looking at an adapted version of the matrix in Figure 9. It has been discussed that high degrees of automation are present on the matrix's right-hand side, but to illustrate the levels of automation, a dashed diagonal line originating from the model's center and extending outward has been included. The red dots signify points in the quadrant with different levels of automation. Hence, we can immediately discern that it is plausible for SOCs in the same quadrant (high automation + high human) to possess different levels of automation (i.e., points 2A and 2B, with the former having lower LOA than the latter). Appendix E provides two scenarios of two different SOCs operating with varying levels of automation.

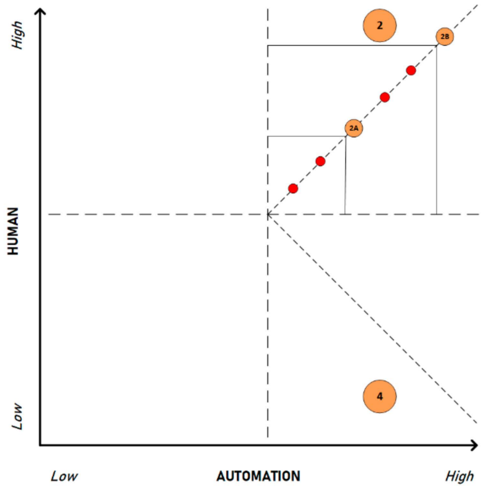

**Figure 9.** SOC levels of automation.

### *4.3. Limitations and Future Research*

Due to the article's brevity and available resources, only four databases for relevant literature were explored. Additionally, we formulated the search terms based on our knowledge of the research domain. Therefore, when sourcing literature, the likelihood of error or missing data exists due to the human factor. Future studies aim to develop an instrument that empirically measures automation bias and complacency among security analysts. This, coupled with observations and interviews in real-world SOCs, will permit SOCs to position themselves on the matrix and work toward optimization.

## **5. Conclusions and Contribution**

This review emphasized the core challenges that security operation center analysts face, necessitating automation integration. Following this, we highlighted the application areas of SOC automation and illustrated the degrees and varying levels to which it can be implemented. The optimal system characteristics, represented in Figure 5, were derived from understanding where and how automation is effectively used in SOCs. The applicability of automation throughout the incident response lifecycle indicates the technological advancement in this domain, moving beyond static incident detection and primarily driven by AI and ML solutions. Coupled with its application areas and the system characteristics that automation should incorporate for a conducive ideal analyst-in-the-loop environment, the human factor implications and sentiment illustrate that, based on its implementation, automation effects in a SOC can vary. Therefore, the SOC Automation Matrix marks this study's primary contribution, representing four quadrants comprising varying human-automation collaboration strategies. The characteristics of each quadrant were delineated, with Q2 being identified as the most desirable. This matrix permits SOCs to position themselves based on current practices and take active steps toward optimal performance. We posit that most SOCs fluctuate between Q1 and Q4 in today's technological landscape due to the plethora of automated tools on the market. However, optimization of these tools appears to be a work in progress. Furthermore, we provided a clear distinction between the degrees and levels of automation, showcasing how the two phenomena interact. Combined with the system characteristics discussed, the matrix allows for a conducive human–automation environment to be established—one that prioritizes the socio-technical design elements. To our knowledge, this appears to be the first human–automation framework within the SOC context. Theoretically, this study works towards developing explicit constructs to measure automation bias and complacency. Together, the matrix and system characteristics assist in mitigating automation bias and complacency.

**Supplementary Materials:** The following supporting information can be downloaded at https://www.mdpi.com/article/10.3390/jcp4030020/s1, A spreadsheet titled Table S1: Scoping Review Literature Overview Per Paper. Table S1 contains a summarized description of all 49 articles included in the final selection. Articles are summarized by the following headers: Problem, Automation Aspect, Automation Aspect Explained, Contribution/Findings, Technology Implemented, Human Factor Sentiment, and Discussed the role of the SOC analyst.

**Author Contributions:** Conceptualization, J.T. and S.F.; methodology, J.T. and S.F.; software, J.T.; validation, J.T. and S.F.; formal analysis, J.T.; investigation, J.T.; resources, J.T. and S.F; data curation, J.T.; writing—original draft preparation, J.T.; writing—review and editing, J.T. and S.F.; visualization, J.T.; supervision, S.F.; project administration, S.F.; All authors have read and agreed to the published version of the manuscript.

**Funding:** This research received no external funding.

**Data Availability Statement:** The original contributions presented in the study are included in the article and Supplementary Materials, and further inquiries can be directed to the corresponding author.

**Acknowledgments:** We would like to acknowledge the University of Tulsa Cyber Fellows initiative.

**Conflicts of Interest:** The authors declare no conflicts of interest.

## Appendix A. Search Strings by Select Database

Table A1 displays the two search strings that were applied to the ScienceDirect database. Each search string relates to the research objectives. The same search string was applied to the other databases, factoring in their semantic requirements.

**Table A1.** ScienceDirect Search Strings.

| Database | Search String | Results |
|---|---|---|
| ScienceDirect (199) | **Secondary Objective #1**<br>Title, abstract, or author-specified keywords (“Security Operations Center” OR “Network Operation Center” OR “Cybersecurity Operation” OR “Cyber Security Operation” OR “Incident Response”) AND (“Automate” OR “Automation” OR “Decision Support” OR “Decision Aid”) | 39 |
| | **Secondary Objective #2**<br>Title, abstract, or author-specified keywords (“Security Operations Center” OR “Network Operations Center” OR “Cybersecurity” OR “Cyber Security”) AND (“SOC Analyst” OR “Analyst” OR “Security Analyst” OR “Human”) AND Find articles with these terms (“Automate” OR “Automation” OR “Decision Support” OR “Decision Aid” OR “Technical Control”) AND (“Complacency” OR “Bias” OR “Trust”) | 160 |
| | Total after removing duplicates: 193 | |

## Appendix B. SOC Challenges Thematic Maps

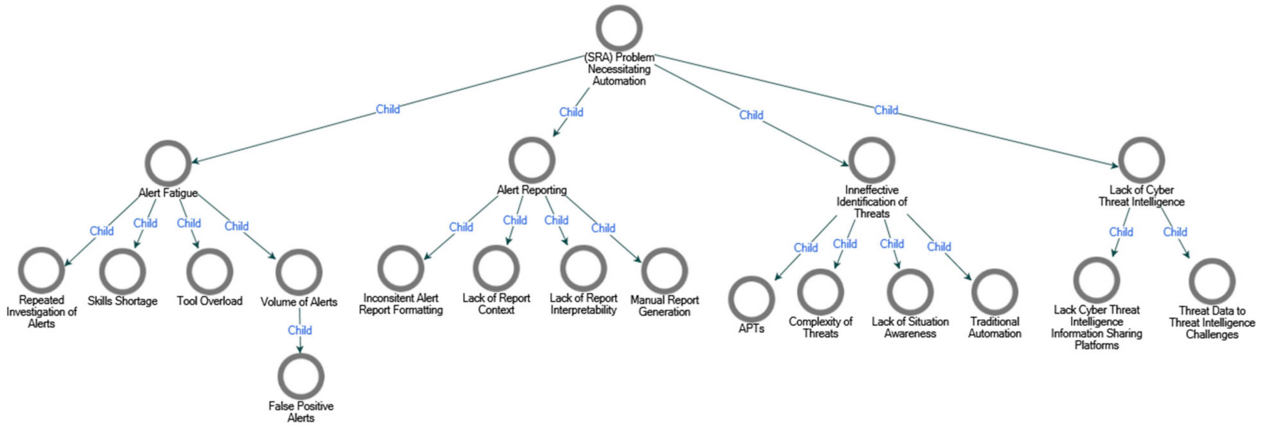
(SRA) Problem Necessitating Automation
Child
Child
Child
Child
Alert Fatigue
Child
Child
Child
Child
Repeated Investigation of Alerts
Skills Shortage
Tool Overload
Volume of Alerts
Child
False Positive Alerts
Alert Reporting
Child
Child
Child
Child
Inconsitent Alert Report Formatting
Lack of Report Context
Lack of Report Interpretability
Manual Report Generation
Inneffective Identification of Threats
Child
Child
Child
Child
APTs
Complexity of Threats
Lack of Situation Awareness
Traditional Automation
Lack of Cyber Threat Intelligence
Child
Child
Lack Cyber Threat Intelligence Information Sharing Platforms
Threat Data to Threat Intelligence Challenges

**Figure A1.** The thematic map above displays the SOC challenges that necessitate automation, broken down into four sub-themes.

## Appendix C. Automation Application Areas Thematic Map

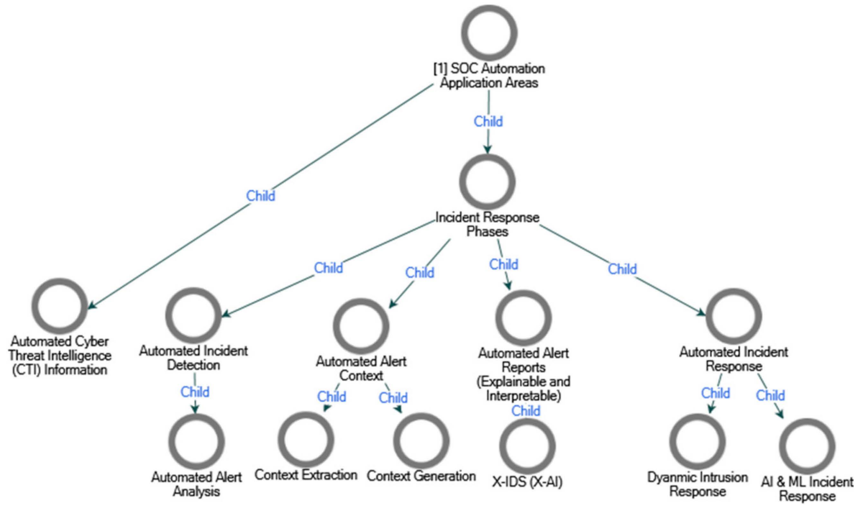
[1] SOC Automation Application Areas
Child
Child
Automated Cyber Threat Intelligence (CTI) Information
Incident Response Phases
Child
Child
Child
Child
Automated Incident Detection
Child
Automated Alert Analysis
Automated Alert Context
Child
Child
Context Extraction
Context Generation
Automated Alert Reports (Explainable and Interpretable)
Child
X-IDS (X-AI)
Automated Incident Response
Child
Child
Dyanmic Intrusion Response
AI & ML Incident Response

**Figure A2.** The thematic map above displays the SOC automation application areas, broken down into two sub-themes and discussed in terms of the incident response lifecycle.

## Appendix D. SOC Automation Restrictive Characteristics

Table A2 displays restrictive characteristics of SOC automation, describes each characteristic, and mentions the study's results that informed it.

**Table A2.** SOC Automation Restrictive Characteristics.

| Characteristic | Details | Results Informing Characteristic |
|---|---|---|
| Black-box automation | Automation lacks transparency in how it operates. | - Alert Reporting Constraints<br>- Automated Alert Reports (Explainable and Interpretable) |
| Automation competes on low false-negative rates | The premise is not to let any alert slip through. This comes at the cost of not fine-tuning sensors, resulting in high alert volumes. | - Ineffective Identification of Threats<br>- Alert Fatigue (Volume of Alerts)<br>- Automated Incident Detection |
| Tool overload | SOCs are inundated with new automated tools instead of optimizing the use of select tools. | - Ineffective Identification of Threats<br>- Alert Fatigue (Tool Overload) |
| Unsupervised machine learning | Tools are not trained based on the expertise of SOC analyst insights but learn independently. | - Automated Incident Response<br>- AI and ML in IR |

## Appendix E. SOC Automation Scenarios

Table A3 delineates the two scenarios shown in Figure 9; representing different levels of SOC automation

**Table A3.** SOC Automation Scenarios.

| SOC | Automated Processes and Levels of Automation |
|---|---|
| SOC 2A | • Automation is likely to be implemented and utilized in the IR lifecycle's detection, prioritization, analysis, and containment phases.<br>• Decision support systems are adopted.<br>• The levels of automation fall between levels 3 and 5.<br>o L3: Automation generates a variety of options that analysts do not have to follow.<br>o L4: Automation selects an action, but analysts approve or disapprove.<br>o L5: Automation selects an action to be taken and implements it if analysts approve. |
| SOC 2B | • Automation is likely to be implemented in the same phases as SOC 2A and the response phase (eradication and recovery procedures).<br>• Decision support systems are adopted.<br>• The levels of automation fall between levels 6 and 8.<br>o L6: Automation selects an action intending to implement it but provides analysts ample time to reverse the action.<br>o L7: Automation performs a task and informs analysts what was performed.<br>o L8: Automation performs a task and informs analysts only if prompted. |

## References

1. Chiba, D.; Akiyama, M.; Otsuki, Y.; Hada, H.; Yagi, T.; Fiebig, T.; Van Eeten, M. DomainPrio: Prioritizing Domain Name Investigations to Improve SOC Efficiency. *IEEE Access* **2022**, *10*, 34352–34368. [CrossRef]
2. Husák, M.; Čermák, M. SoK: Applications and Challenges of Using Recommender Systems in Cybersecurity Incident Handling and Response. In *ARES '22: Proceedings of the 17th International Conference on Availability, Reliability and Security, Vienna Austria, 23–26 August 2022*; Association for Computing Machinery: New York, NY, USA, 2022. [CrossRef]
3. Coro Cybersecurity. 2024 SME Security Workload Impact Report; Coro Cybersecurity. 2023; pp. 1–11. Available online: https://www.coro.net/sme-security-workload-impact-report (accessed on 12 January 2024).
4. Singh, I.L.; Molloy, R.; Parasuraman, R. Individual Differences in Monitoring Failures of Automation. *J. Gen. Psychol.* **1993**, *120*, 357–373. [CrossRef]

5. Skitka, L.J.; Mosier, K.L.; Burdick, M. Does Automation Bias Decision-Making? *Int. J. Hum.-Comput. Stud.* **1999**, *51*, 991–1006. [CrossRef]
6. Brown, P.; Christensen, K.; Schuster, D. An Investigation of Trust in a Cyber Security Tool. *Proc. Hum. Factors Ergon. Soc. Annu. Meet.* **2016**, *60*, 1454–1458. [CrossRef]
7. Butavicius, M.; Parsons, K.; Lillie, M.; McCormac, A.; Pattinson, M.; Calic, D. When believing in technology leads to poor cyber security: Development of a trust in technical controls scale. *Comput. Secur.* **2020**, *98*, 102020. [CrossRef]
8. Cain, A.A. Trust and Complacency in Cyber Security. Master's Thesis, San Jose State University, San Jose, CA, USA, 2016. [CrossRef]
9. Chen, J.; Mishler, S.; Hu, B. Automation Error Type and Methods of Communicating Automation Reliability Affect Trust and Performance: An Empirical Study in the Cyber Domain. *IEEE Trans. Human-Mach. Syst.* **2021**, *51*, 463–473. [CrossRef]
10. Lyn Paul, C.; Blaha, L.M.; Fallon, C.K.; Gonzalez, C.; Gutzwiller, R.S. Opportunities and Challenges for Human-Machine Teaming in Cybersecurity Operations. *Proc. Hum. Factors Ergon. Soc. Annu. Meet.* **2019**, *63*, 442–446. [CrossRef]
11. Ryan, T.J.; Alarcon, G.M.; Walter, C.; Gamble, R.; Jessup, S.A.; Capiola, A.; Pfahler, M.D. Trust in Automated Software Repair: The Effects of Repair Source, Transparency, and Programmer Experience on Perceived Trustworthiness and Trust. In Proceedings of the HCI for Cybersecurity, Privacy and Trust, Orlando, FL, USA, 26–31 July 2019; Moallem, A., Ed.; Springer International Publishing: Berlin/Heidelberg, Germany, 2019; Volume 11594, pp. 452–470. [CrossRef]
12. Bridges, R.A.; Rice, A.E.; Oesch, S.; Nichols, J.A.; Watson, C.; Spakes, K.; Norem, S.; Huettel, M.; Jewell, B.; Weber, B.; et al. Testing SOAR Tools in Use. *Comput. Secur.* **2023**, *129*, 103201. [CrossRef]
13. Altamimi, S.; Altamimi, B.; Côté, D.; Shirmohammadi, S. Toward a Superintelligent Action Recommender for Network Operation Centers Using Reinforcement Learning. *IEEE Access* **2023**, *11*, 20216–20229. [CrossRef]
14. Crowley, C.; Filkins, B.; Pescatore, J. *SANS 2023 SOC Survey*; SANS Analyst Program; SANS Institure: Rockville, MA, USA, 2023; pp. 1–24. Available online: https://www.sans.org/white-papers/2023-sans-soc-survey/ (accessed on 8 February 2024).
15. Agyepong, E.; Cherdantseva, Y.; Reinecke, P.; Burnap, P. A Systematic Method for Measuring the Performance of a Cyber Security Operations Centre Analyst. *Comput. Secur.* **2023**, *124*, 102959. [CrossRef]
16. Shahjee, D.; Ware, N. Integrated Network and Security Operation Center: A Systematic Analysis. *IEEE Access* **2022**, *10*. [CrossRef]
17. Sheridan, T.B.; Parasuraman, R. Human-Automation Interaction. *Rev. Hum. Factors Ergon.* **2005**, *1*, 89–129. [CrossRef]
18. Hauptman, A.I.; Schelble, B.G.; McNeese, N.J.; Madathil, K.C. Adapt and Overcome: Perceptions of Adaptive Autonomous Agents for Human-AI Teaming. *Comput. Hum. Behav.* **2023**, *138*. [CrossRef]
19. Miller, C.A.; Parasuraman, R. Beyond Levels of Automation: An Architecture for More Flexible Human-Automation Collaboration. *Proc. Hum. Factors Ergon. Soc. Annu. Meet.* **2003**, *47*, 182–186. [CrossRef]
20. Applebaum, A.; Johnson, S.; Limiero, S.; Smith, M. Smith Playbook Oriented Cyber Response. In Proceedings of the 2018 National Cyber Summit (NCS), Huntsville, AL, USA, 5 June 2018; pp. 8–15. [CrossRef]
21. Peters, M.D.J.; Marnie, C.; Tricco, A.C.; Pollock, D.; Munn, Z.; Alexander, L.; McInerney, P.; Godfrey, C.M.; Khalil, H. Updated Methodological Guidance for the Conduct of Scoping Reviews. *JBI Evid. Synth.* **2020**, *18*, 2119–2126. [CrossRef] [PubMed]
22. Braun, V.; Clarke, V. Using Thematic Analysis in Psychology. *Qual. Res. Psychol.* **2006**, *3*, 77–101. [CrossRef]
23. Chamkar, S.A.; Maleh, Y.; Gherabi, N. The Human Factor Capabilities in Security Operation Centers (SOC). *EDPACS* **2022**, *66*, 1–14. [CrossRef]
24. Kokulu, F.B.; Soneji, A.; Bao, T.; Shoshitaishvili, Y.; Zhao, Z.; Doupé, A.; Ahn, G.-J. Matched and Mismatched SOCs: A Qualitative Study on Security Operations Center Issues. In *CCS '19; Proceedings of the 2019 ACM SIGSAC Conference on Computer and Communications Security, London, UK, 11–15 November 2019*; Association for Computing Machinery: New York, NY, USA, 2019; pp. 1955–1970. [CrossRef]
25. Vielberth, M.; Bohm, F.; Fichtinger, I.; Pernul, G. Security Operations Center: A Systematic Study and Open Challenges. *IEEE Access* **2020**, *8*, 227756–227779. [CrossRef]
26. Amthor, P.; Fischer, D.; Kühnhauser, W.E.; Stelzer, D. Automated Cyber Threat Sensing and Responding: Integrating Threat Intelligence into Security-Policy-Controlled Systems. In *ARES '19, Proceedings of the 14th International Conference on Availability, Reliability and Security, Canterbury, UK, 26–29 August 2019*; Association for Computing Machinery: New York, NY, USA, 2019. [CrossRef]
27. Yang, W.; Lam, K.-Y. Automated Cyber Threat Intelligence Reports Classification for Early Warning of Cyber Attacks in Next Generation SOC. In *Information and Communications Security*; ICICS 2019. Lecture Notes in Computer Science; Zhou, J., Luo, X., Shen, Q., Xu, Z., Eds.; Springer: Berlin/Heidelberg, Germany, 2020; Volume 11999, pp. 145–164. [CrossRef]
28. Baroni, P.; Cerutti, F.; Fogli, D.; Giacomin, M.; Gringoli, F.; Guida, G.; Sullivan, P. Self-Aware Effective Identification and Response to Viral Cyber Threats. In Proceedings of the 13th International Conference on Cyber Conflict (CyCon), Tallinn, Estonia, 25–28 May 2021; Jancarkova, T., Lindstrom, L., Visky, G., Zotz, P., Eds.; NATO CCD COE Publications. CCD COE: Tallinn, Estonia, 2021; Volume 2021, pp. 353–370. [CrossRef]
29. Strickson, B.; Worsley, C.; Bertram, S. Human-Centered Assessment of Automated Tools for Improved Cyber Situational Awareness. In Proceedings of the 2023 15th International Conference on Cyber Conflict: Meeting Reality (CyCon), Tallinn, Estonia, 30 May–2 June 2023; pp. 273–286. [CrossRef]
30. Islam, C.; Babar, M.A.; Croft, R.; Janicke, H. SmartValidator: A Framework for Automatic Identification and Classification of Cyber Threat Data. *J. Network. Comput. Appl.* **2022**, *202*, 103370. [CrossRef]

31. Basyurt, A.S.; Fromm, J.; Kuehn, P.; Kaufhold, M.-A.; Mirbabaie, M. Help Wanted-Challenges in Data Collection, Analysis and Communication of Cyber Threats in Security Operation Centers. In Proceedings of the 17th International Conference on Wirtschaftsinformatik, Online. 21–23 February 2022. Available online: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85171997510&partnerID=40&md5=30a02b455898c7c2c9d2421d82606470 (accessed on 15 September 2023).
32. Hughes, K.; McLaughlin, K.; Sezer, S. Dynamic Countermeasure Knowledge for Intrusion Response Systems. In Proceedings of the Irish Signals and Systems Conference (ISSC), Letterkenny, Ireland, 11–12 June 2020. [CrossRef]
33. Gupta, N.; Traore, I.; de Quinan, P.M.F. Automated Event Prioritization for Security Operation Center Using Deep Learning. In Proceedings of the 2019 IEEE International Conference on Big Data (Big Data), Los Angeles, CA, USA, 9–12 December 2019; pp. 5864–5872. [CrossRef]
34. Renners, L.; Heine, F.; Kleiner, C.; Rodosek, G.D. Adaptive and Intelligible Prioritization for Network Security Incidents. In Proceedings of the 2019 International Conference on Cyber Security and Protection of Digital Services (Cyber Security), Oxford, UK, 3–4 June 2019; pp. 1–8. [CrossRef]
35. Van der Kleij, R.; Schraagen, J.M.; Cadet, B.; Young, H. Developing Decision Support for Cybersecurity Threat and Incident Managers. *Comput. Secur.* **2022**, *113*, 102535. [CrossRef]
36. Ban, T.; Samuel, N.; Takahashi, T.; Inoue, D. Combat Security Alert Fatigue with AI-Assisted Techniques. In *CSET '21: Proceedings of the 14th Cyber Security Experimentation and Test Workshop, Virtual, 9 August 2021*; Association for Computing Machinery: New York, NY, USA, 2021; Volume 21, pp. 9–16. [CrossRef]
37. Neupane, S.; Ables, J.; Anderson, W.; Mittal, S.; Rahimi, S.; Banicescu, I.; Seale, M. Explainable Intrusion Detection Systems (X-IDS): A Survey of Current Methods, Challenges, and Opportunities. *IEEE Access* **2022**, *10*, 112392–112415. [CrossRef]
38. Goodall, J.R.; Ragan, E.D.; Steed, C.A.; Reed, J.W.; Richardson, D.; Huffer, K.; Bridges, R.; Laska, J. Situ: Identifying and Explaining Suspicious Behavior in Networks. *IEEE Trans. Vis. Comput. Graph.* **2019**, *25*, 204–214. [CrossRef] [PubMed]
39. Ndichu, S.; Ban, T.; Takahashi, T.; Inoue, D. A Machine Learning Approach to Detection of Critical Alerts from Imbalanced Multi-Appliance Threat Alert Logs. In Proceedings of the IEEE International Conference on Big Data (Big Data), Orlando, FL, USA, 15–18 December 2021; pp. 2119–2127. [CrossRef]
40. Ofte, H.J.; Katsikas, S. Understanding Situation Awareness in SOCs, a Systematic Literature Review. *Comput. Secur.* **2023**, *126*, 103069. [CrossRef]
41. Akinrolabu, O.; Agrafiotis, I.; Erola, A. The Challenge of Detecting Sophisticated Attacks: Insights from SOC Analysts. In Proceedings of the 13th International Conference on Availability, Reliability and Security, Hamburg, Germany, 27–30 August 2018; ACM: Hamburg, Germany, 2018; pp. 1–9. [CrossRef]
42. Yen, T.-F.; Oprea, A.; Onarlioglu, K.; Leetham, T.; Robertson, W.; Juels, A.; Kirda, E. Beehive: Large-Scale Log Analysis for Detecting Suspicious Activity in Enterprise Networks. In *ACSAC '13: Proceedings of the 29th Annual Computer Security Applications Conference, New Orleans, LA, USA, 9–13 December 2013*; Association for Computing Machinery: New York, NY, USA, 2013; pp. 199–208. [CrossRef]
43. Zhong, C.; Yen, J.; Lui, P.; Erbacher, R. Learning From Experts' Experience: Toward Automated Cyber Security Data Triage. *IEEE Syst. J.* **2019**, *13*, 603–614. [CrossRef]
44. Alahmadi, B.A.; Axon, L.; Martinovic, I. 99% False Positives: A Qualitative Study of SOC Analysts' Perspectives on Security Alarms. In Proceedings of the 31st Usenix Security Symposium, Boston, MA, USA, 10–12 August 2022; Usenix—The Advanced Computing Systems Association: Berkeley, CA, USA, 2022.
45. Liu, J.; Zhang, R.; Liu, W.; Zhang, Y.; Gu, D.; Tong, M.; Wang, X.; Xue, J.; Wang, H. Context2Vector: Accelerating Security Event Triage via Context Representation Learning. *Inf. Softw. Technol.* **2022**, *146*, 106856. [CrossRef]
46. Gutzwiller, R.S.; Fugate, S.; Sawyer, B.D.; Hancock, P.A. The Human Factors of Cyber Network Defense. In Proceedings of the Human Factors and Ergonomics Society Annual Meeting, Los Angeles, CA, USA, 26–30 October 2015; Volume 59, pp. 322–326. [CrossRef]
47. Chung, M.-H.; Yang, Y.; Wang, L.; Cento, G.; Jerath, K.; Raman, A.; Lie, D.; Chignell, M.H. Implementing Data Exfiltration Defense in Situ: A Survey of Countermeasures and Human Involvement. *ACM Comput. Surv.* **2023**, *55*. [CrossRef]
48. Sworna, Z.T.; Islam, C.; Babar, M.A. APIRO: A Framework for Automated Security Tools API Recommendation. *ACM Trans. Softw. Eng. Methodol.* **2023**, *32*. [CrossRef]
49. Hassan, W.U.; Guo, S.; Li, D.; Chen, Z.; Jee, K.; Li, Z.; Bates, A. NoDoze: Combatting Threat Alert Fatigue with Automated Provenance Triage. In Proceedings of the 2019 Network and Distributed System Security Symposium, San Diego, CA, USA, 24–27 February 2019. [CrossRef]
50. Happa, J.; Agrafiotis, I.; Helmhout, M.; Bashford-Rogers, T.; Goldsmith, M.; Creese, S. Assessing a Decision Support Tool for SOC Analysts. *Digit. Threats: Res. Pract.* **2021**, *2*, 1–35. [CrossRef]
51. Afzaliseresht, N.; Miao, Y.; Michalska, S.; Liu, Q.; Wang, H. From Logs to Stories: Human-Centred Data Mining for Cyber Threat Intelligence. *IEEE Access* **2020**, *8*, 19089–19099. [CrossRef]
52. Kurogome, Y.; Otsuki, Y.; Kawakoya, Y.; Iwamura, M.; Hayashi, S.; Mori, T.; Sen, K. EIGER: Automated IOC Generation for Accurate and Interpretable Endpoint Malware Detection. In *ACSAC '19: Proceedings of the 35th Annual Computer Security Applications Conference, San Juan, PR, USA, 9–13 December 2019*; Association for Computing Machinery: New York, NY, USA, 2019; pp. 687–701. [CrossRef]

53. Dietrich, C.; Krombholz, K.; Borgolte, K.; Fiebig, T. Investigating System Operators' Perspective on Security Misconfigurations. In Proceedings of the ACM SIGSAC Conference on Computer and Communications Security, Toronto, ON, Canada, 15–19 October 2018; ACM: Toronto, ON, Canada, 2018; pp. 1272–1289. [CrossRef]
54. Oprea, A.; Li, Z.; Norris, R.; Bowers, K. MADE: Security Analytics for Enterprise Threat Detection. In *ACSAC'18: Proceedings of the 34th Annual Computer Security Applications Conference, San Juan, PR, USA, 3–7 December 2018*; Association for Computing Machinery: New York, NY, USA, 2018; pp. 124–136. [CrossRef]
55. Kinyua, J.; Awuah, L. Ai/Ml in Security Orchestration, Automation and Response: Future Research Directions. *Intell. Autom. Soft Comp.* **2021**, *28*, 527–545. [CrossRef]
56. Van Ede, T.; Aghakhani, H.; Spahn, N.; Bortolameotti, R.; Cova, M.; Continella, A.; van Steen, M.; Peter, A.; Kruegel, C.; Vigna, G. DEEPCASE: Semi-Supervised Contextual Analysis of Security Events. In Proceedings of the 2022 IEEE Symposium on Security and Privacy (SP), San Francisco, CA, USA, 23–25 May 2022; pp. 522–539. [CrossRef]
57. Chen, C.; Lin, S.C.; Huang, S.C.; Chu, Y.T.; Lei, C.L.; Huang, C.Y. Building Machine Learning-Based Threat Hunting System from Scratch. *Digital Threats* **2022**, *3*, 1–21. [CrossRef]
58. Erola, A.; Agrafiotis, I.; Happa, J.; Goldsmith, M.; Creese, S.; Legg, P.A. RicherPicture: Semi-Automated Cyber Defence Using Context-Aware Data Analytics. In Proceedings of the 2017 International Conference on Cyber Situational Awareness, Data Analytics And Assessment (Cyber SA), London, UK, 19–20 June 2017; pp. 1–8. [CrossRef]
59. Naseer, A.; Naseer, H.; Ahmad, A.; Maynard, S.B.; Masood Siddiqui, A. Real-Time Analytics, Incident Response Process Agility and Enterprise Cybersecurity Performance: A Contingent Resource-Based Analysis. *Int. J. Inf. Manag.* **2021**, *59*, 102334. [CrossRef]
60. Andrade, R.O.; Yoo, S.G. Cognitive Security: A Comprehensive Study of Cognitive Science in Cybersecurity. *J. Inf. Secur. Appl.* **2019**, *48*. [CrossRef]
61. Husák, M.; Sadlek, L.; Špaček, S.; Laštovička, M.; Javorník, M.; Komárková, J. CRUSOE: A Toolset for Cyber Situational Awareness and Decision Support in Incident Handling. *Comput. Secur.* **2022**, *115*, 102609. [CrossRef]
62. Chamberlain, L.B.; Davis, L.E.; Stanley, M.; Gattoni, B.R. Automated Decision Systems for Cybersecurity and Infrastructure Security. In Proceedings of the 2020 IEEE Security and Privacy Workshops (SPW), San Francisco, CA, USA, 18–20 May 2020; pp. 196–201. [CrossRef]
63. González-Granadillo, G.; González-Zarzosa, S.; Diaz, R. Security Information and Event Management (SIEM): Analysis, Trends, and Usage in Critical Infrastructures. *Sensors* **2021**, *21*, 4759. [CrossRef] [PubMed]
64. Demertzis, K.; Tziritas, N.; Kikiras, P.; Sanchez, S.L.; Iliadis, L. The next Generation Cognitive Security Operations Center: Adaptive Analytic Lambda Architecture for Efficient Defense against Adversarial Attacks. *Big Data Cogn. Comput.* **2019**, *3*, 6. [CrossRef]
65. Tilbury, J.; Flowerday, S. The Rationality of Automation Bias in Security Operation Centers. *J. Inf. Syst. Secur.* **2024**, *20*, 87–107.
66. Janssen, C.P.; Donker, S.F.; Brumby, D.P.; Kun, A.L. History and Future of Human-Automation Interaction. *Int. J. Hum.-Comput. Stud.* **2019**, *131*, 99–107. [CrossRef]
67. Haltofová, P.; Štěpánková, P. An Application of the Boston Matrix within Financial Analysis of NGOs. *Procedia-Soc. Behav. Sci.* **2014**, *147*, 56–63. [CrossRef]
68. Du, M.; Li, F.; Zheng, G.; Srikumar, V. DeepLog: Anomaly Detection and Diagnosis from System Logs through Deep Learning. In Proceedings of the 2017 ACM SIGSAC Conference on Computer and Communications Security, Dallas, TX, USA, 30 October–3 November 2017; ACM: New York, NY, USA; pp. 1285–1298. [CrossRef]
69. Ianni, M.; Masciari, E. SCOUT: Security by Computing OUTliers on Activity Logs. *Comput. Secur.* **2023**, *132*, 103355. [CrossRef]
70. Van der Schyff, K.; Flowerday, S.; Lowry, P.B. Information Privacy Behavior in the Use of Facebook Apps: A Personality-Based Vulnerability Assessment. *Heliyon* **2020**, *6*, e04714. [CrossRef] [PubMed]
71. Endsley, M.R.; Kaber, D.B. Level of Automation Effects on Performance, Situation Awareness and Workload in a Dynamic Control Task. *Ergonomics* **1999**, *42*, 462–492. [CrossRef] [PubMed]
72. Tilbury, J.; Flowerday, S. Automation Bias and Complacency in Security Operation Centers. *Computers* **2024**. *Forthcoming*.

**Disclaimer/Publisher's Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.