REVISTA TÓPICOS
ISSN 2965-6672 | Qualis A2

AUTOMAÇÃO E ORQUESTRAÇÃO NA RESPOSTA A INCIDENTES: DA GOVERNANÇA SOAR À CONTENÇÃO E PERÍCIA EM ENDPOINTS COM PYTHON

INCIDENT RESPONSE AUTOMATION AND ORCHESTRATION: FROM SOAR GOVERNANCE TO ENDPOINT CONTAINMENT AND FORENSICS WITH PYTHON

Ciências Exatas e da Terra, Engenharias • 22/11/2026

REGISTRO DOI: 10.70773/revistatopicos/795171074

Caio Cezar Jotta Nogueira[1]

Yago da Costa Jardim Alves Braga[2]

Gabriel Victor Martins Carvalho[3]

Daniel Monteiro de Carvalho[4]

Christiane Martins de Vasconcellos Silveira[5]

Wellington Ávila[6]

RESUMO

Centros de Operações de Segurança (SOCs) operam sob constante sobrecarga: campanhas de *ransomware* e invasões destrutivas concluem etapas críticas em minutos, mas a triagem manual mantém o tempo de permanência do invasor (*dwell time*) mediano em cerca de 10 dias. Intervenções improvisadas nas máquinas monitoradas também costumam destruir vestígios voláteis na memória RAM, inviabilizando investigações forenses e de resposta a incidentes (DFIR). Este trabalho analisa como o emprego de automação e orquestração atua nesse cenário por meio de uma revisão sistemática da literatura (RSL), seguindo o protocolo PRISMA 2020 e as orientações de Kitchenham e Charters para Engenharia de Software. O levantamento reuniu 17 estudos (15 primários e 2 secundários) publicados entre 2019 e 2026 nas bases ACM Digital Library, IEEE Xplore, Google Scholar e Revista Tópicos. A avaliação de qualidade dos artigos (QA1 a QA5) alcançou índice de concordância quase perfeito entre avaliadores (κ = 0,812) e nota média de 3,09 pontos em 5,00 possíveis. Como resultados centrais, o artigo propõe uma taxonomia em quatro eixos operacionais e detalha uma arquitetura híbrida de dois níveis, conectando motores SOAR baseados na especificação aberta OASIS CACAO 2.0 a rotinas de contenção em Python executadas nos *endpoints*. A literatura revisada comprova reduções expressivas de tempo, incluindo queda de 77% no MTTR em bancada aberta (Neuber, 2025) e de 61% no tempo de detecção com geração dinâmica de casos (Fayyad, 2025). Por outro lado, o receio de indisponibilidade por falsos alarmes, erros factuais em modelos de linguagem (42% de inconsistências sem supervisão em Kramer et al., 2025) e exigências legais da cadeia de custódia impõem a manutenção de supervisão humana (*human-in-the-loop*). O estudo conclui apresentando uma política dual de triagem e quatro diretrizes práticas de engenharia para

implementação segura em ambientes corporativos.
Palavras-chave: Automação de Segurança; SOAR; Resposta a Incidentes; Python; Eficiência Operacional.

ABSTRACT

Security Operations Centers (SOCs) operate under constant strain: ransomware campaigns and destructive attacks complete critical stages in minutes, but manual triage keeps the median dwell time at around 10 days. Ad-hoc interventions on monitored machines also often destroy volatile traces in RAM, rendering digital forensics and incident response (DFIR) investigations impossible. This study analyzes how the use of automation and orchestration addresses this scenario through a systematic literature review (SLR), following the PRISMA 2020 protocol and the guidelines by Kitchenham and Charters for Software Engineering. The survey included 17 studies (15 primary and 2 secondary) published between 2019 and 2026 in the ACM Digital Library, IEEE Xplore, Google Scholar, and Revista Tópicos databases. The quality assessment of the articles (QA1 to QA5) achieved a near-perfect agreement index among reviewers (κ = 0.812) and an average score of 3.09 out of a possible 5.00. As key findings, the article proposes a taxonomy based on four operational axes and details a two-tier hybrid architecture, connecting SOAR engines based on the open OASIS CACAO 2.0 specification to Python containment routines executed at the endpoints. The reviewed literature demonstrates significant time reductions, including a 77% decrease in MTTR in an open testbed (Neuber, 2025) and a 61% reduction in detection time with dynamic case generation (Fayyad, 2025). On the other hand, concerns about unavailability due to false alarms, factual errors in language models (42% of unsupervised inconsistencies in Kramer et al., 2025), and legal requirements for the chain of custody necessitate the maintenance of human supervision

(human-in-the-loop). The study concludes by presenting a dual triage policy and four practical engineering guidelines for safe implementation in corporate environments.
Keywords: Security Automation; SOAR; Incident Response; Python; Operational Efficiency.

1. INTRODUÇÃO

A rotina dos Centros de Operações de Segurança (SOCs) tem sido tensionada pelo aumento na frequência e na velocidade dos ataques cibernéticos. Equipes técnicas precisam filtrar volumes massivos de alertas enquanto invasores conseguem comprometer ambientes internos e iniciar a exfiltração de dados em poucos minutos (Islam; Babar; Nepal, 2019; Gurabi et al., 2024). Relatórios recentes do setor confirmam essa disparidade: o Mandiant M-Trends (Mandiant, 2024) documenta um tempo mediano de permanência de invasores (*dwell time*) de 10 dias, enquanto o relatório da Verizon (Verizon, 2024) destaca a velocidade de infecção em incidentes de *ransomware*. Na prática, procedimentos manuais não padronizados em estações de trabalho costumam alterar o estado da memória física dos sistemas, destruindo provas cruciais para a perícia digital e a resposta a incidentes (DFIR) (Kuforiji, 2025).

Diante desse gargalo, o mercado passou a investir em plataformas de orquestração, automação e resposta de segurança (SOAR), cujo foco é unificar a telemetria, enriquecer alertas com dados de inteligência e executar fluxos predefinidos (*playbooks*) (Islam; Babar; Nepal, 2019; Tsirakis et al., 2025). Diretamente nos *endpoints*, por sua vez, a execução de rotinas em linguagens de fácil integração com chamadas de sistema, especialmente Python, viabiliza o bloqueio seletivo de conexões de rede, o encerramento seguro de processos e

a coleta de artefatos probatórios (Muthusamy et al., 2024; Rodolà, 2024; Stelly; Roussev, 2019).

A literatura científica, no entanto, frequentemente trata esses dois níveis de forma isolada. Trabalhos voltados a SOAR tendem a priorizar aspectos conceituais ou permanecem atrelados a produtos proprietários (Tsirakis et al., 2025). Em contrapartida, estudos focados no *host* limitam-se a provas de conceito locais sem vínculo com padrões abertos de governança (Muthusamy et al., 2024). Além disso, há poucas revisões sistemáticas que meçam o impacto quantitativo da automação sobre tempos operacionais, como o tempo médio de detecção (MTTD, do inglês *Mean Time to Detect*) e o tempo médio de resposta (MTTR, *Mean Time to Respond*), e discutam o atrito entre acelerar o isolamento da máquina e preservar a cadeia de custódia segundo normas reconhecidas, a exemplo da RFC 3227 (Brezinski; Killalea, 2002), da ISO/IEC 27037 (ISO/IEC, 2012) e dos arts. 158-A a 158-F do Código de Processo Penal brasileiro (Brasil, 2019).

Para responder a essas lacunas, este estudo apresenta uma revisão sistemática da literatura estruturada segundo o protocolo PRISMA 2020 (Page et al., 2021) e as diretrizes de Kitchenham e Charters (2007). A partir de um *corpus* de 17 estudos submetidos a avaliação metodológica formal, o trabalho estabelece cinco contribuições: uma arquitetura híbrida de dois níveis; uma taxonomia em quatro eixos para classificação de soluções; a síntese crítica de indicadores operacionais obtidos na literatura; a análise dos riscos de modelos generativos, falsos positivos e contaminação probatória, acompanhada de uma política dual de triagem (*Fast-Path* versus *Forensic-Path*); e quatro diretrizes práticas de engenharia para SOCs corporativos.

Em relação ao escopo, este trabalho extrai diretamente dos 17 estudos da revisão os parâmetros de governança, especificações CACAO 2.0 e o estado da arte de modelos de aprendizado aplicados a SOCs. Já a descrição das rotinas locais em Python (gerenciamento de processos com *psutil*, isolamento por firewall e chamadas a coletores de memória em *kernel*) atua como uma proposta de engenharia baseada na documentação técnica oficial das ferramentas, integrando o controle central do SOC à atuação no dispositivo físico.

2. FUNDAMENTAÇÃO TEÓRICA

2.1. Orquestração de Segurança, SOAR e Padronização de Playbooks

Plataformas SOAR atuam como centros de coordenação no SOC: integram a recepção de alertas, agregam fontes de inteligência de ameaças (CTI) e automatizam etapas de resposta por meio de sequências estruturadas (*playbooks*) (Islam; Babar; Nepal, 2019). Por muito tempo, a falta de formatos interoperáveis dificultou o compartilhamento desses procedimentos entre organizações que utilizam ferramentas distintas. Para mitigar esse problema, o consórcio OASIS (2023) estabeleceu o padrão aberto CACAO 2.0 (*Collaborative Automated Course of Action Operations*, *Committee Specification 01*, nov. 2023). O modelo adota sintaxe JSON para especificar rotinas de mitigação, versionamento, assinaturas digitais e pontos de decisão com intervenção humana obrigatória (*human-in-the-loop*) (Tsirakis et al., 2025; Paduraru; Dumitru; Stefanescu, 2025). Além de rotinas estáticas, pesquisas recentes exploram modelos de aprendizado para sugerir ações em tempo real (Kremer et al., 2023), classificar incidentes com algoritmos como *Random*

Forest (Neuber, 2025) e resumir grandes volumes de *logs* com técnicas de recuperação aumentada (RAG) (Cadet et al., 2026).

2.2. Mecanismos Locais de Execução e o Papel dos Scripts em Python

Se o motor SOAR planeja a resposta na central de operações, as ações concretas de contenção precisam ser executadas diretamente na máquina atingida (Kuforiji, 2025). No *endpoint*, *scripts* desenvolvidos em Python assumem tarefas operacionais cruciais. Como rotinas comuns em espaço de usuário (*user-space*) não interceptam pacotes na camada de rede, a automação comumente aciona firewalls nativos do sistema operacional (*netsh advfirewall* no Windows, *iptables* ou *nftables* no Linux). Essa abordagem fecha conexões externas e internas suspeitas, mas resguarda a porta de comunicação remota com o SOC para permitir o controle do ativo. Na manipulação de processos, a biblioteca *psutil* possibilita mapear identificadores de processos (PIDs, *Process Identifiers*), verificar portas de rede abertas e emitir comandos de pausa (SIGSTOP) ou encerramento (SIGKILL / TerminateProcess). Como esclarece Rodolà (2024), o *psutil* consulta as interfaces padrão do sistema sem inserir drivers em nível de *kernel* (*ring-0*). Isso significa que a biblioteca não acessa a memória física bruta e não identifica ameaças ocultas por técnicas avançadas de *rootkit*, como a manipulação direta de objetos do *kernel* (DKOM, *Direct Kernel Object Manipulation*), demandando a atuação combinada com ferramentas de detecção e resposta em *endpoints* (EDR, *Endpoint Detection and Response*) (Kuforiji, 2025). Para extração de memória RAM em perícias, faz-se necessário disparar binários assinados de *kernel*, como WinPmem no Windows (Cohen, 2024), *Acquire Volatile Memory for Linux* (AVML) (Microsoft, 2024) ou *Linux Memory Extractor* (LiME) (Sylve,

2024) no Linux. O *script* em Python gerencia essa captura e calcula o *hash* SHA-256 imediato da imagem gerada via biblioteca *hashlib*, atrelando-o a um carimbo de tempo certificado (RFC 3161; Adams et al., 2001) para atender aos padrões de custódia da prova digital.

3. METODOLOGIA

A revisão sistemática foi executada seguindo as diretrizes do PRISMA 2020 (Page et al., 2021) e o protocolo de Engenharia de Software formulado por Kitchenham e Charters (2007).

3.1. Questões de Pesquisa (QPs)

Três questões de pesquisa nortearam a coleta de dados e a análise comparativa dos artigos, conforme listado na Tabela 1.

Tabela 1. Questões de Pesquisa (QPs) da Revisão Sistemática

ID	Foco Temático	Questão de Pesquisa Formulada	Seção
QP1	Mecanismos e Arquiteturas	Quais mecanismos técnicos, padrões abertos de *playbooks* e arquiteturas de orquestração têm sido empregados para coordenar a automação em SOCs e a execução local em *endpoints*?	Seções 4.1 e 4.2
QP2	Eficiência e Métricas	Qual o impacto quantitativo real da automação sobre métricas de desempenho operacional (MTTD, MTTR e tempo de contenção) em comparação com processos manuais?	Seção 4.3

QP3	Desafios e Governança	Quais os principais desafios técnicos, riscos operacionais (falsos alarmes, alucinações de IA) e tensões de preservação pericial (cadeia de custódia) na automação?	Seções 4.4 e 4.5

Fonte: Elaborada pelos autores (2026).

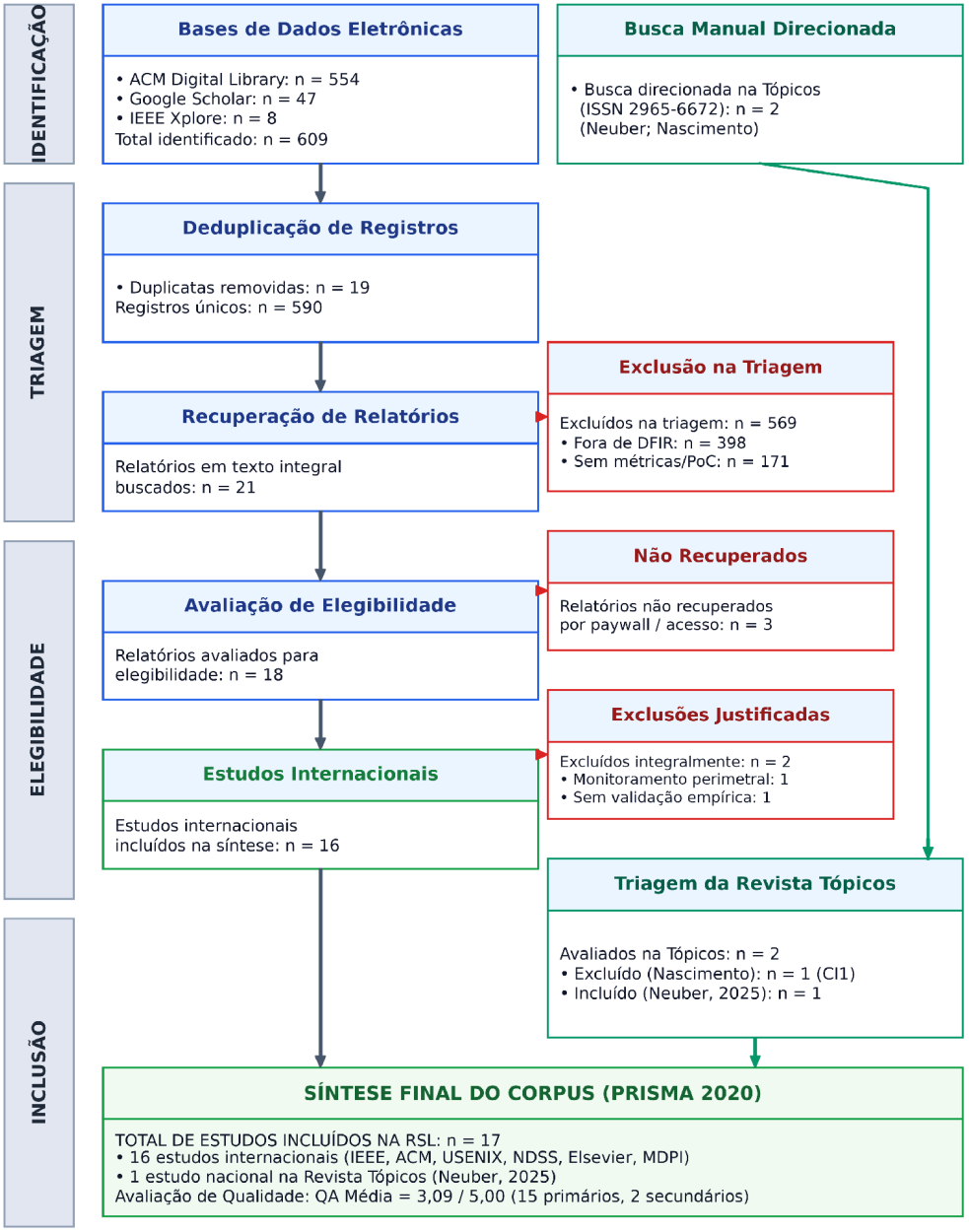
3.2. Estratégia de Busca, Deduplicação e Protocolo PRISMA 2020

A consulta às bases ocorreu em 02 de novembro de 2026 nos repositórios ACM Digital Library (554 registros), Google Scholar (47 registros) e IEEE Xplore (8 registros), somando 609 trabalhos. Em paralelo, realizou-se uma pesquisa complementar nos arquivos da Revista Tópicos (ISSN 2965-6672) para mapear experimentos desenvolvidos no Brasil com foco em ferramentas de código aberto para SOCs. Empregou-se a seguinte expressão booleana: ("incident response" OR "security orchestration" OR "SOAR") AND ("automation" OR "Python" OR "scripting" OR "playbook") AND ("containment" OR "quarantine" OR "evidence collection" OR "DFIR"). Na ACM e na IEEE, a pesquisa varreu título, resumo e palavras-chave entre 2019 e 2026; Na ACM; utilizou-se a coleção.*ACM Guide to Computing Literature*, que também indexa trabalhos de outras editoras e simpósios de computação, como *NDSS Symposium*, *USENIX SOUPS* e revistas da Elsevier (*Computers & Security* e *Procedia CS*). No Google Scholar, os termos nucleares de resposta a incidentes foram delimitados ao título (*intitle:* / *allintitle:*), combinados aos descritores de automação e perícia nos metadados indexados (filtro temporal de 2019 a 2026, excluindo citações e patentes, avaliando os 50 primeiros resultados por relevância). Esse

procedimento retornou 47 artigos de periódicos abertos e repositórios de *preprints* (*MDPI*, *arXiv*, *IJTMH*, *JDACM* e *IJRASET*).

Após a coleta dos 609 registros, o descarte de duplicatas removeu 19 itens (restando 590 únicos). A leitura de títulos e resumos excluiu 569 trabalhos fora do escopo do estudo (398 não abordavam DFIR/SOAR e 171 não apresentavam métricas ou implementação prática), deixando 21 relatórios para leitura integral. Desses, 3 não puderam ser acessados por restrições de *paywall* (CE4). Dos 18 artigos analisados por completo, 2 foram desconsiderados (1 por tratar apenas de monitoramento de tráfego de rede e 1 por se restringir a discussão teórica sem validação prática). Na busca complementar da Revista Tópicos (2 trabalhos examinados), desconsiderou-se o estudo de Nascimento et al. (2026), voltado a fluxos gerais de DevOps com n8n sem aplicação em segurança, e incluiu-se o trabalho de Neuber (2025), que testou em laboratório um arranjo com Wazuh, TheHive, a plataforma de compartilhamento de ameaças *Malware Information Sharing Platform* (MISP) e aprendizado de máquina. Dessa forma, o *corpus* consolidou-se em 17 estudos (15 primários e 2 secundários), ilustrados no diagrama PRISMA (Figura 1).

Figura 1. Fluxograma do processo de seleção dos estudos segundo o protocolo PRISMA 2020



Fonte: Elaborada pelos autores com base nas diretrizes do protocolo PRISMA 2020 (Page et al., 2021).

3.3. Critérios de Elegibilidade e Análise de Sensibilidade

A seleção de artigos baseou-se em quatro Critérios de Inclusão (CI1 a CI4) e quatro Critérios de Exclusão (CE1 a CE4), descritos na Tabela 2.

Tabela 2. Critérios de Elegibilidade (Inclusão e Exclusão)

Categoria	ID	Descrição do Critério de Seleção Aplicado

Inclusão	CI1	Estudo focado em automação, orquestração técnica (SOAR, *playbooks*) ou rotinas em *endpoints* para resposta a incidentes e DFIR.
Inclusão	CI2	Apresentar descrição explícita de arquitetura técnica, fluxo de *playbooks*, rotinas de *script* (em particular Python) ou integração de ferramentas.
Inclusão	CI3	Conter avaliação empírica quantitativa (MTTD, MTTR, contenção, *recall*/acurácia), prova de conceito experimental ou levantamento formal de requisitos.
Inclusão	CI4	Publicado entre 2019 e 2026 em inglês ou português em anais revisados por pares, periódicos indexados ou *preprints* consolidados.
Exclusão	CE1	Artigo focado exclusivamente em detecção perimetral sem fluxos de resposta, quarentena ou contenção.
Exclusão	CE2	Trabalho sem arquitetura técnica verificável, código/*playbooks* ou metodologia clara de implementação.
Exclusão	CE3	Artigos de opinião editorial, resenhas curtas (< 3 páginas), *slides* ou estudos sem revisão científica formal.
Exclusão	CE4	Relatório com texto integral inacessível por *paywall* contratual ou restrições de confidencialidade comercial.

Fonte: Elaborada pelos autores (2026).

A seleção cobriu arquiteturas consolidadas de orquestração, rotinas locais de *script* e aplicações emergentes de aprendizado de máquina e de modelos de linguagem de grande e pequeno porte (LLMs, *Large Language Models*, e SLMs, *Small Language Models*) na

montagem de *playbooks*. Para testar a estabilidade da pontuação média de qualidade (QA), foi realizada uma análise de sensibilidade removendo os dois trabalhos secundários do *corpus* (Islam; Babar; Nepal, 2019, QA = 3,0; Husák; Čermák, 2022, QA = 2,5). Sem esses mapeamentos, os 15 estudos primários somam 47,0 pontos em 75,0 possíveis, o que eleva levemente a média de 3,09 para 3,13 pontos (62,7% do valor máximo, variação de +0,04 ponto). A proporção entre as faixas de rigor permaneceu a mesma (1 estudo em Alto Rigor, 11 em Rigor Moderado e 3 em Rigor Exploratório), assegurando que a inclusão das duas revisões não distorceu as conclusões técnicas do trabalho.

3.4. Avaliação de Qualidade Metodológica dos Estudos (Quality Assessment - QA)

Para avaliar o rigor metodológico dos artigos, construiu-se um formulário estruturado com base nas recomendações de Kitchenham e Charters (2007) e Dybå e Dingsøyr (2008). O instrumento contempla cinco critérios (QA1 a QA5), especificados no Quadro 1. Dois avaliadores pontuaram cada estudo de forma independente, utilizando notas 1,0 (critério plenamente atendido), 0,5 (parcialmente atendido) e 0,0 (critério não atendido ou ausente), totalizando até 5,0 pontos por artigo (máximo de 85,0 pontos no conjunto dos 17 estudos).

Quadro 1. Questões de Avaliação de Qualidade Metodológica (QA1 a QA5) e Regras de Pontuação

ID	Dimensão Avaliada	Questão de Qualidade Formulada	Regra de Pontuação (1,0 | 0,5 | 0,0)

QA1	Alinhamento e Rigor Metodológico	O estudo define objetivos claros de automação/orquestração em resposta a incidentes e explicita seu método de pesquisa?	1,0 = Objetivo e método explícitos; 0,5 = Parcialmente descritos; 0,0 = Ausentes.
QA2	Especificação Arquitetural e Técnica	O trabalho detalha arquitetura verificável, esquema formal de *playbooks* (ex.: CACAO 2.0) ou implementação de *scripts*?	1,0 = Arquitetura/esquema formal; 0,5 = Modelo conceitual/parcial; 0,0 = Sem detalhe técnico.
QA3	Validação Empírica e Métricas	Há avaliação empírica (experimento, *testbed* ou *survey*) com métricas quantitativas (MTTD, MTTR, acurácia ou *recall*)?	1,0 = Experimento amplo com métricas; 0,5 = PoC local, bancada ou *survey*; 0,0 = Sem validação.
QA4	Governança, HITL e Riscos Operacionais	O estudo analisa governança operacional, supervisão humana (*human-in-the-loop*), falsos alarmes ou alucinações de IA?	1,0 = Foco central em governança; 0,5 = Abordagem parcial de HITL; 0,0 = Não aborda riscos.
QA5	Preservação Forense e Cadeia de Custódia	A solução contempla coleta/preservação de evidências voláteis,	1,0 = Coleta/custódia direta no *host*;

		hashing ou rastreabilidade forense?	0,5 = Telemetria auditável/assinatura; 0,0 = Sem forense.

Fonte: Elaborado pelos autores com base em Kitchenham e Charters (2007) e Dybå e Dingsøyr (2008).

O cruzamento das avaliações nas 85 células da matriz (17 estudos multiplicados pelos 5 critérios) revelou 75 notas idênticas (concordância observada Po = 75/85 = 0,8824) e 10 discordâncias pontuais que foram resolvidas em consenso conjunto (concordância esperada Pe = 0,3741). Esse alinhamento resultou em um Coeficiente Kappa de Cohen de $\kappa = 0,812$, caracterizando concordância quase perfeita. No cômputo geral, os 17 trabalhos somaram 52,5 pontos dos 85,0 possíveis, equivalente a uma média de 3,09 pontos (61,8% do total). Como o critério inicial QA1 serviu como pré-requisito de entrada (recebendo 1,0 em todos os artigos incluídos), a pontuação nas quatro frentes técnicas (QA2 a QA5) somou 35,5 de 68,0 pontos (média de 0,52 por item). Na divisão por faixas (Tabela 3), 1 trabalho (5,9%) atingiu Alto Rigor (nota ≥ 4,0: Nitz et al., 2025); 12 trabalhos (70,6%) ficaram na faixa de Rigor Moderado (3,0 a 3,5 pontos); e 4 trabalhos (23,5%) enquadraram-se em Rigor Exploratório (nota 2,5: Neuber, 2025; Husák; Čermák, 2022; Kuforiji, 2025; Muthusamy et al., 2024). Pontuações menores em implementação de código e bancada experimental nas revisões de Islam, Babar e Nepal (2019) e Husák e Čermák (2022) refletem a natureza teórica de mapeamento secundário desses textos, e não fragilidade metodológica.

Tabela 3. Matriz de Avaliação de Qualidade Metodológica dos Estudos Incluídos (QA)

Estudo	Veículo / Origem	QA1	QA2	QA3	QA4
Hammar; Alpcan; Lupu (2026)	NDSS Symposium	1,0	1,0	1,0	0,5
Tsirakis et al. (2025)	Elsevier Computers &	1,0	1,0	0,5	0,5

⚠ Esta tabela possui muitas colunas e foi cortada para impressão. Para visualizá-la completa, acesse o artigo original em: https://revistatopicos.com.br/artigos/automacao-e-orquestracao-na-resposta-a-incidentes-da-governanca-soar-a-contencao-e-pericia-em-endpoints-com-python?noblockage

Fonte: Elaborada pelos autores com base em Kitchenham e Charters (2007) e Dybå e Dingsøyr (2008).

3.5. Síntese e Caracterização do Corpus Selecionado

A Tabela 4 sintetiza o delineamento metodológico, o foco tecnológico e a contribuição central de cada um dos 17 estudos que compõem o *corpus* da revisão.

Tabela 4. Estudos Selecionados na RSL (15 Primários e 2 Secundários)

Estudo	Tipologia Metodológica	Foco Tecnológico Principal	Contribuição Principal para a RSL
Cadet et al. (2026)	Experimento primário com	RAG sobre *logs* (Active	*Recall* médio de 89%–94% em *malware* (100%

	RAG/LLM (CAIS '26)	Directory e *malware*)	vs. 38%–50% sem RAG no *Fake Auth*) e até 96% em ataques Active Directory.
Falowo; Abdo (2026)	*Survey* empírico primário (194 profissionais)	Prontidão de SOCs, barreiras e *AI Trust*	87% rejeitam decisões autônomas sem intervenção humana em cenários críticos; 83% avaliam que riscos de falsos alarmes superam os benefícios.
Fayyad (2025)	Experimento quantitativo (200 cenários simulados)	Geração Dinâmica de Casos e MITRE ATT&CK	Redução de 61% no tempo médio de detecção (46 min para 18 min), ganho na acurácia (72% para 91%) e > 50% menos ações manuais.
Gurabi et al. (2024)	Levantamento de requisitos (ACM DTRAP)	Engenharia de Requisitos de *Playbooks*	Elicitação de requisitos para automação com *playbooks* legíveis por máquina e tratamento formal de exceções.
Hammar; Alpcan; Lupu (2026)	Experimento primário com LLMs leves (NDSS)	*Lightweight* LLM, *Fine-Tuning* e RAG	Planejamento de resposta via LLM leve com limite probabilístico de alucinação e recuperação até 22% mais rápida.
Husák; Čermák (2022)	Estudo secundário (SoK – ACM ARES)	Sistemas de Recomendaçã o em Resposta	Sistematização do estado da arte em recomendação de ações, barreiras de confiança e adoção em SOCs.
Islam; Babar; Nepal (2019)	Estudo secundário	SOAR, Orquestração	Mapeamento taxonômico de SOAR e

	(Mapeamento – ACM CSUR)	e Integração	identificação da heterogeneidade de APIs como gargalo estrutural.
Kramer et al. (2025)	Estudo empírico (18 analistas Google, 50 casos)	LLMs na Sumarização de Incidentes	42% de erros factuais em resumos automáticos; preferência por modelo copiloto humano-IA em 77% dos casos.
Kremer et al. (2023)	Experimento primário com *Deep Learning*	IC-SECURE: Recomendador de *Playbooks*	Rede neural profunda para recomendação e composição adaptativa de etapas de *playbooks* em SOAR.
Kuforiji (2025)	Proposta conceitual e arquitetural (JDACM)	Automação DFIR com IA, EDR e SOAR	*Framework* integrando contenção via EDR, *snapshots* de memória volátil e cadeia de custódia auditável.
Lee (2023)	Experimento primário (ACM ICSCA)	Enriquecimento de Artefatos Forenses (eLog)	Mecanismo automatizado correlacionando artefatos forenses a plataformas CTI/SOAR (MISP, TheHive, Cortex) com 94,4% de adequação.
Muthusamy et al. (2024)	Prova de conceito primária (IJRASET)	Python no SOC e Forense (Scapy, Pytsk, Volatility)	Automação em Python para monitoramento, isolamento de *host* em ataques de *ransomware* e análise forense de disco e memória.
Neuber (2025)	Experimento em bancada	Wazuh, TheHive, Cortex, MISP,	Redução média de 77% no MTTR em 3 cenários simulados e contenção

	(Revista Tópicos)	*Random Forest*	pontual em 12 s (isolamento físico em ~8 s).
Nitz et al. (2025)	Arquitetura interorganizacional (ACM DTRAP)	Resposta Colaborativa e Privacidade	*Framework* interorganizacional para resposta colaborativa com sanitização e anonimização de IoCs em *Threat Intelligence*.
Paduraru; Dumitru; Stefanescu (2025)	Experimento primário com LLMs (Elsevier)	Geração de *Playbooks* CACAO 2.0 via LLMs	*Pipeline* com LLMs compactos gerando *playbooks* OASIS CACAO 2.0 a partir de dados estruturados de incidentes.
Stelly; Roussev (2019)	Experimento em linguagem DFIR (ACM ARES)	Linguagem de Domínio Específico (DSL)	DSL declarativa (Nugget) integrando consulta de *logs* (Splunk) e extração forense remota em *hosts* (GRR).
Tsirakis et al. (2025)	Experimento ontológico (Computers & Security)	Ontologia e Validação de *Playbooks* CACAO	Ontologia formal e sistema de conhecimento para validação de integridade e interoperabilidade de CACAO 2.0.

Fonte: Elaborada pelos autores a partir dos dados extraídos do *corpus* da RSL (2026).

3.6. Ameaças à Validade

A validade de construção ancorou-se em normas e guias consolidados pela indústria, incluindo NIST SP 800-61r3 (NIST, 2025),

OASIS CACAO 2.0 (OASIS, 2023), RFC 3227 (Brezinski; Killalea, 2002) e ISO/IEC 27037 (ISO/IEC, 2012). Na dimensão QA5, aplicou-se o princípio de equivalência funcional entre procedimentos internacionais e os artigos de custódia da legislação brasileira, evitando distorções regionais na avaliação dos estudos. No que se refere à validade interna, convém destacar a diversidade do perfil editorial das fontes: o levantamento abrangeu tanto simpósios de ponta e periódicos com fator de impacto consagrado (como NDSS, USENIX SOUPS, ACM CSUR e Elsevier *Computers & Security*) quanto publicações abertas, bancadas de teste e repositórios de artigos preliminares (*IJRASET*, *IJTMH*, *JDACM*, *Revista Tópicos* e *arXiv*). Por essa razão, os dados numéricos de estudos exploratórios devem ser interpretados como demonstrações de viabilidade e não como referências absolutas de mercado. Outras restrições incluem a filtragem de termos no título no Google Scholar para limitar o ruído nos resultados, a impossibilidade de baixar 3 textos retidos por *paywall* (CE4) e a restrição da busca manual nacional à Revista Tópicos. Por fim, a validade externa é limitada pelo tamanho da amostra (17 artigos) e pelo fato de a maioria dos experimentos ter ocorrido em laboratórios virtuais controlados, o que exige cautela antes de transpor diretamente essas taxas para operações corporativas reais.

4. RESULTADOS E DISCUSSÕES

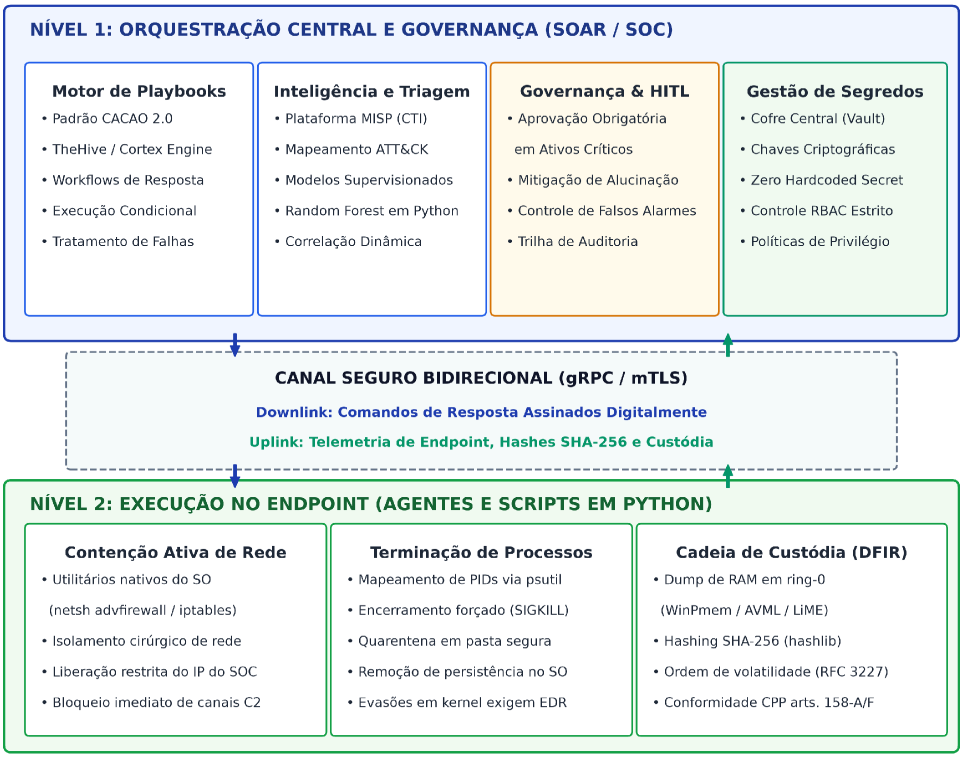
4.1. Mecanismos e Arquiteturas Técnicas de Orquestração (QP1)

A análise do *corpus* mostra que 7 dos 17 estudos (41,2%) tratam de modelos analíticos, levantamento de requisitos ou fatores humanos de adoção (Hammar; Alpcan; Lupu, 2026; Cadet et al., 2026; Falowo; Abdo, 2026; Gurabi et al., 2024; Kramer et al., 2025; Islam; Babar;

Nepal, 2019; Husák; Čermák, 2022). Já entre os 10 artigos que implementam soluções práticas, observam-se três formatos predominantes: (i) orquestração puramente centralizada em plataformas SOAR ou nuvem (3 artigos, 17,6%: Tsirakis et al., 2025; Paduraru; Dumitru; Stefanescu, 2025; Kremer et al., 2023); (ii) atuação direta na máquina local (3 artigos, 17,6%: Lee, 2023; Muthusamy et al., 2024; Kuforiji, 2025); e (iii) modelos híbridos integrados, formato mais comum com 4 artigos (23,5%: Neuber, 2025; Fayyad, 2025; Nitz et al., 2025; Stelly; Roussev, 2019).

A partir dessa distribuição, propõe-se a Arquitetura Híbrida de Dois Níveis esquematizada na Figura 2. O Nível 1 assume a governança central do SOC: interpreta *playbooks* em OASIS CACAO 2.0 (usando motores como TheHive e Cortex), correlaciona indicadores de inteligência (MISP), realiza triagem assistida e gerencia credenciais em cofres seguros (HashiCorp Vault), exigindo validação de analistas humanos antes de ações que possam interromper serviços. O envio de instruções aos sistemas monitorados ocorre por um canal protegido via segurança mútua da camada de transporte (mTLS, *Mutual Transport Layer Security*) e chamadas de procedimento remoto (gRPC, *gRPC Remote Procedure Calls*), transportando comandos assinados digitalmente e recebendo de volta registros de telemetria, metadados e *hashes* de evidências. O Nível 2 executa as ações localmente no dispositivo via Python: aplica regras de firewall (*netsh advfirewall* ou *iptables*), controla processos por meio do *psutil* e aciona ferramentas de *kernel* (WinPmem, AVML ou LiME) para despejo de memória volátil, assegurando a rastreabilidade da prova pericial.

Figura 2. Proposta de Arquitetura Híbrida de Dois Níveis para Resposta a Incidentes



Fonte: Elaborada pelos autores (2026).

4.2. Proposição de Taxonomia Multidimensional para Resposta Automatizada a Incidentes

A carência de esquemas conceituais integrando o plano central do SOC à resposta em *endpoints* (Islam; Babar; Nepal, 2019) motivou a criação de uma taxonomia em quatro eixos operacionais. O Eixo 1 (Gatilho e Triagem) distingue o modo de ativação da resposta, separando fluxos reativos (disparados por telemetria de estações ou alertas de segurança), proativos (orientados por inteligência CTI e matriz MITRE ATT&CK) e federados (cooperação entre diferentes organizações). O Eixo 2 (Decisão e Raciocínio) categoriza a lógica aplicada, abrangendo desde *playbooks* determinísticos e aprendizado supervisionado até sistemas de recomendação em *deep learning* e modelos generativos (LLMs/SLMs). O Eixo 3 (Execução Técnica) organiza o ponto em que as ações ocorrem: macro-orquestração no servidor, micro-execução local no dispositivo ou modelo híbrido acoplado. Por fim, o Eixo 4 (Governança e DFIR) abrange o nível de supervisão e custódia da evidência, variando de

automação assistida (*human-in-the-loop*) à validação ontológica estrita e extração probatória pericial. A classificação dos 17 estudos seguiu o mesmo protocolo de consenso independente da revisão metodológica (κ = 0,812) e está disposta na Tabela 5.

Tabela 5. Matriz de Classificação Taxonômica dos Estudos do Corpus (n = 17)

Estudo	**Eixo 1: Gatilho e Triagem**	**Eixo 2: Decisão e Raciocínio**	**Eixo 3: Execução Técnica**	**Eixo 4: Governança e DFIR**
Cadet et al. (2026)	Reativo / Telemetria de *Endpoint*	Generativo / LLM e SLM	N/A (Metodológico / *Logs*)	Supervisionado / HITL
Falowo; Abdo (2026)	Reativo / Telemetria de SOC	*Survey* / Prontidão e *AI Trust*	N/A (*Survey* Empírico)	Supervisionado / HITL
Fayyad (2025)	Proativo / CTI e ATT&CK	Aprendizado Supervisionado	Híbrido Integrado	Supervisionado / HITL
Gurabi et al. (2024)	Reativo / Telemetria Local	Determinístico / Requisitos	N/A (Requisitos)	Supervisionado / HITL
Hammar; Alpcan; Lupu (2026)	Reativo / Telemetria de *Endpoint*	Generativo / LLM e SLM	N/A (Metodológico / *Logs*)	Supervisionado / SLM Probabilístico
Husák; Čermák (2022)	Reativo / Telemetria Local	Sistemas de Recomendação	N/A (SoK Secundário)	Supervisionado / HITL
Islam; Babar; Nepal (2019)	Reativo / Telemetria Local	Determinístico / *Playbooks*	N/A (Mapeamento)	Supervisionado / Básico

Kramer et al. (2025)	Reativo / Telemetria Local	Generativo / Modelos LLM	N/A (Sumarização)	Supervisionado / HITL
Kremer et al. (2023)	Reativo / Telemetria Local	*Deep Learning* / Recomendação	Macro-Orquestração	Supervisionado / HITL
Kuforiji (2025)	Reativo / Telemetria Local	Determinístico / Políticas	Micro-Execução no *Endpoint*	Custódia Pericial / Forense
Lee (2023)	Proativo / CTI e Inteligência	Determinístico / Enriquecimento CTI	Micro-Execução no *Endpoint*	Custódia Pericial / Forense
Muthusamy et al. (2024)	Reativo / Telemetria de *Endpoint*	Determinístico / *Playbooks*	Micro-Execução no *Endpoint*	Custódia Pericial / Forense
Neuber (2025)	Reativo / Telemetria de *Endpoint*	Aprendizado Supervisionado / ML	Híbrido Integrado	Autônomo (HITL para alto impacto)
Nitz et al. (2025)	Federado / Interorganizacional	Determinístico / *Playbooks*	Híbrido Integrado	Supervisionado / HITL
Paduraru; Dumitru; Stefanescu (2025)	Reativo / Dados Estruturados de Incidentes	Generativo / LLM e CACAO	Macro-Orquestração	Supervisionado / Validação
Stelly; Roussev (2019)	Reativo / Telemetria Local	Determinístico / DSL DFIR	Híbrido Integrado	Custódia Pericial / Forense

Tsirakis et al. (2025)	Proativo / CTI e Inteligência	Determinístico / CACAO 2.0	Macro-Orquestração	Supervisionado / Formal

Fonte: Elaborada pelos autores (2026).

4.3. Impacto na Eficiência Operacional e Ponderação de Métricas (QP2)

Os indicadores quantitativos de desempenho extraídos dos estudos empíricos do *corpus* encontram-se sintetizados na Tabela 6.

Tabela 6. Síntese Comparativa de Evidências Empíricas Confirmadas no Corpus

Estudo / Veículo	**Parâmetro Mensurado**	**Baseline Manual**	**Resultado Automatizado**	**Impacto Técnico Confirmado**
Fayyad (2025) [IJTMH]	Tempo Médio de Detecção (MDT)	46 min (manual)	18 min (dinâmico)	Redução de 61% no tempo em 200 cenários simulados
Fayyad (2025) [IJTMH]	Acurácia de Resposta (RA)	72% (manual)	91% (automatizada)	Ganho absoluto de +19 p.p. (ganho relativo de +26,4% calculado / 26% no artigo)
Fayyad (2025) [IJTMH]	Carga de Trabalho (AW)	13 ações manuais/caso	6 ações manuais/caso	Redução de 54% (> 50% de esforço manual repetitivo)

Neuber (2025) [Revista Tópicos]	Tempo Médio de Resposta (MTTR)	27,7 min (média manual)	6,0 min (automatizado)	Redução média de 77% (cenários) / 78,3% (média absoluta)
Neuber (2025) [Revista Tópicos]	Ciclo de Resposta (*Ransomware*)	15 min (900 s estimados)	12 segundos (Wazuh + TheHive)	Compressão calculada de 98,7% (97% no estudo original; isolamento em 8 s)
Cadet et al. (2026) [CAIS '26]	*Recall* Semântico em *Logs*	38% a 50% (sem RAG)	100% (*Fake Auth*); 89%–94% médio	Até 100% no *Fake Auth* (vs. 38%–50% sem RAG) e 96% em AD
Kramer et al. (2025) [SOUPS]	Erro Factual IA vs. Preferência	*Baseline* analistas Google	42% erros; 77% preferência humano-IA	Necessidade de HITL; preferência por coprodução humano-IA (77%)

Fonte: Compilada pelos autores a partir dos dados primários dos estudos revisados (2026).

Os dados quantitativos extraídos dividem-se em dois cenários: simulação em larga escala voltada a SOCs corporativos e testes em bancada de laboratório. No primeiro grupo, Fayyad (2025) avaliou a geração dinâmica de casos com base na matriz MITRE ATT&CK em 200 cenários simulados (em contexto corporativo associado à Saudi Aramco). O autor reportou uma redução de 61% no tempo médio de detecção (de 46 para 18 minutos), aumento da acurácia de resposta

de 72% para 91% (representando ganho absoluto de 19 pontos percentuais e ganho relativo de 26,4% sobre a base manual, arredondado para 26% no artigo) e diminuição de 54% nas intervenções manuais por caso (de 13 para 6 ações). Em bancada experimental, Neuber (2025) mediu o impacto de uma pilha de ferramentas abertas (Wazuh, TheHive, Cortex, MISP e classificador *Random Forest* em Python) frente a três vetores de ataque, cujos tempos médios de resposta (MTTR) constam na Tabela 7.

Tabela 7. Tempos Médios de Resposta no Experimento de Neuber (2025)

Tipologia de Incidente Avaliada	**Tempo Manual de Referência**	**Tempo Automatizado Integrado**	**Redução Relativa Alcançada**
Infecção por *Ransomware*	42 minutos	8 minutos	Redução de 81% no tempo de tratamento
Vazamento de Credenciais	26 minutos	6 minutos	Redução de 77% no tempo de tratamento
Escaneamento Interno de Portas	15 minutos	4 minutos	Redução de 73% no tempo de tratamento
Média Geral dos Incidentes	27,7 minutos	6,0 minutos	Redução Média de 77% (cenários) / 78,3% (absoluta)

Fonte: Reproduzida a partir de Neuber (2025, Tabela 3, Revista Tópicos). A redução de 77% corresponde à média simples das taxas relativas dos três cenários (81%, 77% e 73%); sobre as médias absolutas (27,7 min para 6,0 min), a redução atinge 78,3%.

Além das médias de MTTR apresentadas na Tabela 7 (em que o incidente de *ransomware* demandou 8 minutos para tratamento completo), Neuber (2025) conduziu um teste pontual de contenção automatizada concluído em 12 segundos, a partir da identificação do ataque pelo classificador em Python, consulta ao MISP e execução da resposta ativa do Wazuh, atingindo o corte de tráfego do *host* em 8 segundos (compressão calculada de 98,7% perante a estimativa manual de 15 minutos, indicada como 97% pelo autor). Contudo, esse valor pontual deve ser ponderado considerando o perfil laboratorial do estudo (QA = 2,5): como o ensaio ocorreu em ambiente virtualizado isolado sem tráfego de produção concorrente e sem extração preliminar de RAM, o caso ilustra o choque direto entre a velocidade do corte de rede e a perda irreparável de provas voláteis. Nota-se ainda que, embora métricas de tempo médio de contenção (MTTC, *Mean Time to Contain*) sejam populares na indústria, estudos científicos costumam integrá-las ao cálculo amplo de MTTR.

4.4. Desafios Técnicos, Barreiras de Integração e Governança (QP3)

A implantação de automações em SOCs esbarra em desafios práticos que vão além do desenvolvimento de *scripts*. A heterogeneidade de APIs proprietárias e sistemas legados segue como um gargalo histórico (Islam; Babar; Nepal, 2019), sendo mitigada gradativamente por especificações abertas como o OASIS CACAO 2.0 (Gurabi et al., 2024; Tsirakis et al., 2025). A essa complexidade técnica soma-se a aversão das equipes a falsos positivos: na pesquisa empírica de Falowo e Abdo (2026), com 194 analistas de segurança, 87% dos participantes rejeitaram a tomada de decisões totalmente autônomas em situações críticas, e 83%

avaliaram que os riscos operacionais de alarmes falsos superam os benefícios imediatos da automação, receio fundamentado no impacto de paradas acidentais de sistemas em produção (Nitz et al., 2025).

Em relação aos modelos de inteligência artificial generativa, os achados científicos desaconselham a delegação de tarefas sem supervisão humana estrita. O experimento de Kramer et al. (2025) com profissionais da Google constatou que 42% dos resumos de incidentes gerados por LLMs continham imprecisões factuais, levando 77% dos especialistas a optarem por fluxos assistidos sob controle *human-in-the-loop*. Para reduzir essa instabilidade, Hammar, Alpcan e Lupu (2026) demonstraram que modelos compactos (*small language models*), combinados com ajuste fino e técnicas de RAG, conseguem estipular limites probabilísticos contra alucinações e acelerar a recuperação em até 22%. Ademais, quando a resposta exige troca federada de indicadores (IoCs) entre instituições parceiras, regras de privacidade exigem a sanitização automática de dados pessoais, em linha com o Regulamento Geral sobre a Proteção de Dados (União Europeia, 2016; Nitz et al., 2025).

4.5. Coleta e Preservação de Evidências Digitais e a Política Dual

O conflito operacional mais agudo na resposta a incidentes reside na escolha entre conter imediatamente a ameaça ou preservar as evidências digitais. O isolamento instantâneo do dispositivo interrompe a movimentação lateral e o roubo de dados, mas contraria a ordem de volatilidade da RFC 3227 (Brezinski; Killalea, 2002), segundo a qual se deve registrar conexões ativas e tabelas de roteamento antes de modificar o estado da máquina. O corte de rede derruba conexões com servidores de comando e controle (C2),

enquanto a finalização forçada de processos via SIGKILL destrói chaves criptográficas residentes em memória. No ordenamento brasileiro, muito embora os arts. 158-A a 158-F do Código de Processo Penal (Lei nº 13.964/2019; Brasil, 2019) rejam a persecução penal pública, suas exigências de integridade e registro cronológico funcionam como referência essencial para que SOCs corporativos respaldem auditorias e subsidiem eventuais medidas judiciais. Assim, a garantia da prova exige calcular a síntese SHA-256 com aposição de carimbo de tempo conforme a RFC 3161 (Adams et al., 2001) e transferir os artefatos para armazenamento imutável do tipo WORM.

A opção de priorizar o corte de rede em ataques agressivos é respaldada pelo impacto computacional da coleta: a extração de 16 GB a 64 GB de memória RAM por drivers de *kernel* pode levar de 3 a 12 minutos e sobrecarregar o barramento de entrada e saída (I/O) em 20% a 50%, tempo suficiente para que um *ransomware* atinja servidores adjacentes. Para conciliar essa dicotomia, estabelece-se uma Política Dual de Triagem: (i) no fluxo *Fast-Path*, destinado a incidentes de alta destrutividade (*ransomware* e *wipers*), a prioridade é a contenção rápida, realizando um registro preliminar de processos e portas, suspendendo os binários suspeitos com *psutil* (Rodolà, 2024) e aplicando regras de firewall que preservem o canal de controle do SOC, transferindo a coleta de memória para a fase posterior; e (ii) no fluxo *Forensic-Path*, voltado a intrusões silenciosas ou espionagem corporativa, a extração probatória tem preferência, disparando coletores em *kernel* (WinPmem, AVML ou LiME) para despejo integral da RAM antes de qualquer alteração nas interfaces de rede ou processos, em conformidade com as normas periciais (RFC 3227, ISO/IEC 27037 e art. 158-B do CPP).

4.6. Diretrizes de Engenharia para Implementação Segura em SOCs

Com base nos achados empíricos e nas restrições legais analisadas, apresentam-se quatro diretrizes práticas de engenharia para a adoção segura de automação em SOCs:

1. **Desacoplamento arquitetural entre governança e execução:** separar o plano de controle central (estruturado em especificações abertas OASIS CACAO 2.0 sobre plataformas SOAR) do plano de atuação local no *endpoint* implementado por microrrotinas modulares em Python, facilitando a auditoria e reduzindo a dependência de fornecedor (Tsirakis et al., 2025; Muthusamy et al., 2024).

2. **Supervisão humana obrigatória (*human-in-the-loop*) e proteção de ativos críticos:** restringir ações autônomas potencialmente disruptivas por meio de aprovações de analistas e listas de exclusão (*allowlists*) rigorosas para infraestruturas essenciais, reduzindo o risco de indisponibilidade por falsos alarmes e erros factuais de modelos generativos (Falowo; Abdo, 2026; Kramer et al., 2025; Hammar; Alpcan; Lupu, 2026).

3. **Orquestração dinâmica sob Política Dual (*Fast-Path* vs. *Forensic-Path*)**: parametrizar os *playbooks* para alternar condicionalmente entre o isolamento imediato de rede em ataques destrutivos (*Fast-Path*) e a aquisição prévia de memória volátil em nível de *kernel* em investigações sensíveis (*Forensic-Path*), conciliando redução de MTTR e preservação

probatória (Brezinski; Killalea, 2002; Neuber, 2025; Kuforiji, 2025).

4. **Cadeia de custódia automatizada e imutável:** incorporar aos *scripts* de coleta o cálculo imediato de síntese SHA-256 (*hashlib*), a vinculação de carimbo de tempo certificado (RFC 3161; Adams et al., 2001), o transporte autenticado via mTLS e o depósito em armazenamento WORM com sanitização de dados sensíveis quando aplicável (ISO/IEC, 2012; Brasil, 2019; Nitz et al., 2025).

5. CONCLUSÃO

A pesquisa apresentou uma revisão sistemática da literatura sobre a automação e orquestração de resposta a incidentes, analisando 17 estudos (15 primários e 2 secundários) publicados entre 2019 e 2026. A avaliação formal de qualidade metodológica registrou concordância quase perfeita entre os revisores (κ = 0,812) e nota média de 3,09 em 5,00 pontos, apontando a prevalência de estudos em faixas moderadas e exploratórias.

Os dados compilados confirmam que a automação reduz drasticamente os tempos de resposta, com destaque para a queda de 77% no MTTR em bancada laboratorial integrada (Neuber, 2025) e a redução de 61% no tempo de detecção em cenários simulados (Fayyad, 2025). Em contrapartida, a literatura demonstra que a substituição completa de operadores humanos ainda não é segura: analistas apontam riscos elevados de indisponibilidade por falsos alarmes (Falowo; Abdo, 2026; Nitz et al., 2025), modelos de linguagem desassistidos exibem até 42% de erros factuais em

resumos de incidentes (Kramer et al., 2025) e o corte abrupto de conexões pode quebrar a cadeia de custódia da evidência pericial.

Como contribuições práticas para a Engenharia de Software e a Segurança da Informação, o trabalho entregou uma taxonomia em quatro eixos, uma arquitetura híbrida de dois níveis (aliando a governança em CACAO 2.0/SOAR com a execução em Python nos *endpoints*), uma política dual de triagem e quatro diretrizes de engenharia para projetos em SOCs. Para pesquisas futuras, sugere-se testar a interoperabilidade de *playbooks* CACAO 2.0 em redes corporativas heterogêneas, isolar a medição do tempo de contenção (MTTC) frente à sobrecarga de extração de memória RAM e avaliar o uso de modelos de linguagem compactos dotados de mecanismos formais de mitigação de alucinação.

REFERÊNCIAS BIBLIOGRÁFICAS

504ENSICS. **LiME (Linux Memory Extractor): Open source kernel module for volatile memory acquisition.** 2024. Disponível em: https://github.com/504ensicsLabs/LiME. Acesso em: 02 nov. 2026.

ADAMS, C. et al. **Internet X.509 Public Key Infrastructure Time-Stamp Protocol (TSP).** RFC 3161, Internet Engineering Task Force (IETF), ago. 2001. Disponível em: https://www.rfc-editor.org/rfc/rfc3161. Acesso em: 02 nov. 2026.

BRASIL. Decreto-Lei nº 3.689, de 3 de outubro de 1941. **Código de Processo Penal.** [Artigos 158-A a 158-F inseridos pela Lei nº 13.964, de 24 de dezembro de 2019]. Diário Oficial da União, Brasília, DF, 2019.

BREZINSKI, D.; KILLALEA, T. **Guidelines for Evidence Collection and Archiving.** RFC 3227, Internet Engineering Task Force (IETF), fev.

2002. DOI: 10.17487/RFC3227.

CADET, X. et al. Retrieval-Augmented LLMs for Security Incident Analysis. In: **ACM Conference on AI and Agentic Systems (CAIS '26),** May 26–29, 2026, San Jose, CA, USA. New York: ACM, 2026. p. 1-21. DOI: 10.1145/3786335.3813136.

COHEN, M. **WinPmem: Open-source physical memory acquisition driver for Windows.** Velocidex Rekall Project, 2024. Disponível em: https://github.com/Velocidex/WinPmem. Acesso em: 02 nov. 2026.

DYBÅ, T.; DINGSØYR, T. Empirical studies of agile software development: A systematic review. **Information and Software Technology,** Elsevier, v. 50, n. 9-10, p. 833-859, 2008. DOI: 10.1016/j.infsof.2008.01.006.

FALOWO, O. I.; ABDO, J. B. Empirical Study on Automation, AI Trust, and Framework Readiness in Cybersecurity Incident Response. **Algorithms,** MDPI, v. 19, n. 1, art. 62, p. 1-22, 2026. DOI: 10.3390/a19010062.

FAYYAD, D. Enhancing Incident Response Efficiency Through Automated Dynamic Threat Case Generation. **International Journal of Technology, Management and Humanities (IJTMH),** v. 11, n. 4, p. 1-12, 2025. DOI: 10.21590/ijtmh.11.04.04.

GURABI, M. A. et al. Requirements for Playbook-Assisted Cyber Incident Response, Reporting and Automation. **ACM Digital Threats: Research and Practice (DTRAP),** v. 5, n. 3, art. 34, p. 1-11, 2024. DOI: 10.1145/3688810.

HAMMAR, K.; ALPCAN, T.; LUPU, E. C. Incident Response Planning Using a Lightweight Large Language Model with Reduced Hallucination. In: **Network and Distributed System Security (NDSS) Symposium 2026.** San Diego, CA: Internet Society, fev. 2026. p. 1-18. DOI: 10.14722/ndss.2026.240358.

HUSÁK, M.; ČERMÁK, M. SoK: Applications and Challenges of using Recommender Systems in Cybersecurity Incident Handling and Response. In: **17th International Conference on Availability, Reliability and Security (ARES 2022),** Vienna, Austria. New York: ACM, 2022. p. 1-10. DOI: 10.1145/3538969.3538981.

ISLAM, C.; BABAR, M. A.; NEPAL, S. A Multi-Vocal Review of Security Orchestration. **ACM Computing Surveys,** v. 52, n. 2, art. 37, p. 1-45, abr. 2019. DOI: 10.1145/3305268.

ISO/IEC. **ISO/IEC 27037:2012: Information technology – Security techniques – Guidelines for identification, collection, acquisition and preservation of digital evidence.** Genebra: International Organization for Standardization, 2012.

KITCHENHAM, B.; CHARTERS, S. **Guidelines for performing Systematic Literature Reviews in Software Engineering.** Technical Report EBSE 2007-001, Keele University and Durham University, 2007.

KRAMER, D. et al. Integrating Large Language Models into Security Incident Response. In: **21st USENIX Symposium on Usable Privacy and Security (SOUPS).** USENIX Association, 2025. p. 133-148.

KREMER, R. et al. IC-SECURE: Intelligent System for Assisting Security Experts in Generating Playbooks for Automated Incident

Response. **arXiv preprint** arXiv:2311.03825, p. 1-16, 2023. DOI: 10.48550/arXiv.2311.03825.

KUFORIJI, J. Digital Forensics and Incident Response (DFIR) Automation: Leveraging AI to Accelerate Breach Investigation, Evidence Collection, and Cyberattack Mitigation. **Journal of Data Analysis and Critical Management (JDACM),** v. 1, n. 4, p. 1-19, 2025. DOI: 10.64235/tsvfvz27.

LEE, H.-W. Analysis of Digital Forensic Artifacts Data Enrichment Mechanism for Cyber Threat Intelligence. In: **2023 12th International Conference on Software and Computer Applications (ICSCA 2023),** Kuantan, Malaysia. New York: ACM, 2023. p. 192-199. DOI: 10.1145/3587828.3587857.

MANDIANT. **M-Trends 2024: Special Report.** Google Cloud, 2024. Disponível em: https://www.mandiant.com/m-trends. Acesso em: 02 nov. 2026.

MICROSOFT. **Acquire Volatile Memory for Linux (AVML).** Microsoft Open Source, 2024. Disponível em: https://github.com/microsoft/avml. Acesso em: 02 nov. 2026.

MUTHUSAMY, P. et al. Python-Based Security Operations Center (SOC) and Forensics Analysis for Incident Cyber Threats. **International Journal for Research in Applied Science and Engineering Technology (IJRASET),** v. 12, n. 4, p. 2591-2596, abr. 2024. DOI: 10.22214/ijraset.2024.60403.

NASCIMENTO, C. E. P. do et al. Automação de Processos DevOps Utilizando N8N em Infraestrutura Computacional. **Revista Tópicos,** v. 3, n. 28, p. 1-13, 2026.

NEUBER, D. Automação de Defesa Cibernética com Inteligência Artificial e SOAR: Um Framework Aberto para Resposta a Incidentes em Tempo Real. **Revista Tópicos,** v. 3, n. 26, p. 1-14, 2025. DOI: 10.5281/zenodo.17420035. Disponível em: https://revistatopicos.com.br/artigos/automacao-de-defesa-cibernetica-com-inteligencia-artificial-e-soar-um-framework-aberto-para-resposta-a-incidentes-em-tempo-real. Acesso em: 02 nov. 2026.

NIST. **Incident Response Recommendations and Considerations for Cybersecurity Risk Management: A CSF 2.0 Community Profile.** NIST Special Publication 800-61 Revision 3, National Institute of Standards and Technology, Gaithersburg, MD, abr. 2025. DOI: 10.6028/NIST.SP.800-61r3.

NITZ, L. et al. On Collaboration and Automation in the Context of Threat Detection and Response with Privacy-Preserving Features. **ACM Digital Threats: Research and Practice (DTRAP),** v. 6, n. 1, p. 1-36, 2025. DOI: 10.1145/3707651.

OASIS. **Collaborative Automated Course of Action Operations (CACAO) for Cyber Security Version 2.0.** Edited by Bret Jordan and Allan Thomson. OASIS Committee Specification 01, 27 nov. 2023. Disponível em: https://docs.oasis-open.org/cacao/security-playbooks/v2.0/cs01/security-playbooks-v2.0-cs01.html. Acesso em: 02 nov. 2026.

PADURARU, C.; DUMITRU, B.; STEFANESCU, A. Automated Generation of Cybersecurity Response Playbooks via Large Language Models. **Procedia Computer Science,** Elsevier, v. 270, p. 2987-2996, 2025. DOI: 10.1016/j.procs.2025.09.423.

PAGE, M. J. et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. **BMJ,** v. 372, n. 71, p. 1-9, 2021. DOI: 10.1136/bmj.n71.

RODOLÀ, G. **psutil: Cross-platform process and system monitoring in Python.** Versão 5.9.8, 2024. Disponível em: https://psutil.readthedocs.io. Acesso em: 02 nov. 2026.

STELLY, C.; ROUSSEV, V. Language-based integration of digital forensics and incident response. In: **14th International Conference on Availability, Reliability and Security (ARES '19),** Canterbury, United Kingdom. New York: ACM, 2019. p. 1-9. DOI: 10.1145/3339252.3339278.

SYLVE, J. **LiME (Linux Memory Extractor): A Loadable Kernel Module (LKM) for volatile memory acquisition.** 2024. Disponível em: https://github.com/jtsylve/LiME. Acesso em: 02 nov. 2026.

TSIRAKIS, O. et al. Operationalizing Cybersecurity Knowledge: Design, Implementation & Evaluation of a Knowledge Management System for CACAO Playbooks. **Computers & Security,** Elsevier, v. 148, art. 104163, p. 1-18, 2025. DOI: 10.1016/j.cose.2024.104163.

UNIÃO EUROPEIA. Regulamento (UE) 2016/679 do Parlamento Europeu e do Conselho, de 27 de abril de 2016 (Regulamento Geral sobre a Proteção de Dados – GDPR). **Jornal Oficial da União Europeia,** L 119, p. 1-88, 2016.

VERIZON. **2024 Data Breach Investigations Report (DBIR).** Verizon Business, 2024. p. 24. Disponível em: https://www.verizon.com/business/resources/reports/dbir/. Acesso em: 02 nov. 2026.

[1] Discente do Curso de Bacharel em Engenharia de Software da Universidade de Vassouras, Campus Maricá. Av. Roberto Silveira, 1954, Flamengo, Maricá, RJ, CEP 24903-815. ORCID: https://orcid.org/0009-0004-2650-8571. E-mail: acesse o artigo original para visualizar o e-mail.

[2] Discente do Curso de Bacharel em Engenharia de Software da Universidade de Vassouras, Campus Maricá. Av. Roberto Silveira, 1954, Flamengo, Maricá, RJ, CEP 24903-815. ORCID: https://orcid.org/0009-0007-8473-3955. E-mail: acesse o artigo original para visualizar o e-mail.

[3] Discente do Curso de Bacharel em Engenharia de Software da Universidade de Vassouras, Campus Maricá. Av. Roberto Silveira, 1954, Flamengo, Maricá, RJ, CEP 24903-815. ORCID: https://orcid.org/0009-0007-8356-6948. E-mail: acesse o artigo original para visualizar o e-mail.

[4] Discente do Curso de Bacharel em Engenharia de Software da Universidade de Vassouras, Campus Maricá. Av. Roberto Silveira, 1954, Flamengo, Maricá, RJ, CEP 24903-815. ORCID: https://orcid.org/0009-0004-9146-2012. E-mail: acesse o artigo original para visualizar o e-mail.

[5] Doutora em Bioquímica pela Universidade Federal do Rio de Janeiro. Docente do Curso de Bacharel em Engenharia de Software da Universidade de Vassouras, Campus Maricá. Av. Roberto Silveira, 1954, Flamengo, Maricá, RJ, CEP 24903-815. ORCID: https://orcid.org/0000-0003-3825-7242. E-mail: acesse o artigo original para visualizar o e-mail.

[6] Mestre em Gestão do Trabalho pela Universidade Santa Úrsula. Docente do Curso de Bacharel em Engenharia de Software da Universidade de Vassouras, Campus Maricá. Av. Roberto Silveira, 1954, Flamengo, Maricá, RJ, CEP 24903-815. ORCID: https://orcid.org/0000-0002-9377-5993. E-mail: acesse o artigo original para visualizar o e-mail.