DISSERTAÇÃO DE MESTRADO PROFISSIONAL

PROPOSTA DE UM FRAMEWORK PARA ENRIQUECIMENTO DE INTELIGÊNCIA DE AMEAÇAS CIBERNÉTICAS UTILIZANDO INFORMAÇÕES A PARTIR DE INTELIGÊNCIA DE FONTES ABERTAS

Carlos Eduardo de Sousa

Programa de Pós-Graduação Profissional em Engenharia Elétrica

DEPARTAMENTO DE ENGENHARIA ELÉTRICA
FACULDADE DE TECNOLOGIA
UNIVERSIDADE DE BRASÍLIA

UNIVERSIDADE DE BRASÍLIA
Faculdade de Tecnologia

DISSERTAÇÃO DE MESTRADO PROFISSIONAL

PROPOSTA DE UM FRAMEWORK PARA ENRIQUECIMENTO DE INTELIGÊNCIA DE AMEAÇAS CIBERNÉTICAS UTILIZANDO INFORMAÇÕES A PARTIR DE INTELIGÊNCIA DE FONTES ABERTAS

Carlos Eduardo de Sousa

Dissertação de Mestrado Profissional submetida ao Departamento de Engenharia Elétrica como requisito parcial para obtenção do grau de Mestre em Engenharia Elétrica

Banca Examinadora

Prof. João José Costa Gondim, Dr., FT/UnB _____
Orientador

Prof. Laerte Peotta de Melo, Dr., FT/UnB _____
Examinador interno

Prof. Dino Macedo Amaral, Dr., Banco do Brasil _____
Examinador externo

Prof. Georges Daniel Amvame Nze, Dr., FT/UnB _____
Suplente

FICHA CATALOGRÁFICA

SOUSA, CARLOS E.
PROPOSTA DE UM FRAMEWORK PARA ENRIQUECIMENTO DE INTELIGÊNCIA DE AMEAÇAS CIBERNÉTICAS UTILIZANDO INFORMAÇÕES A PARTIR DE INTELIGÊNCIA DE FONTES ABERTAS [Distrito Federal] 2025.
xvi, 58 p., 210 x 297 mm (ENE/FT/UnB, Mestre, Engenharia Elétrica, 2025).
Dissertação de Mestrado Profissional - Universidade de Brasília, Faculdade de Tecnologia.
Departamento de Engenharia Elétrica

1. Resiliência Cibernética	2. Inteligência de Ameaças Cibernéticas
3. Enriquecimento	4. Inteligência de Fontes Abertas
I. ENE/FT/UnB	II. Título (série)

PUBLICAÇÃO: PPEE.MP.094

REFERÊNCIA BIBLIOGRÁFICA

SOUSA, C. E. (2025). *PROPOSTA DE UM FRAMEWORK PARA ENRIQUECIMENTO DE INTELIGÊNCIA DE AMEAÇAS CIBERNÉTICAS UTILIZANDO INFORMAÇÕES A PARTIR DE INTELIGÊNCIA DE FONTES ABERTAS* . Dissertação de Mestrado Profissional, Departamento de Engenharia Elétrica, Universidade de Brasília, Brasília, DF, 58 p.

CESSÃO DE DIREITOS

AUTOR:
TÍTULO: PROPOSTA DE UM FRAMEWORK PARA ENRIQUECIMENTO DE INTELIGÊNCIA DE AMEAÇAS CIBERNÉTICAS UTILIZANDO INFORMAÇÕES A PARTIR DE INTELIGÊNCIA DE FONTES ABERTAS .
GRAU: Mestre em Engenharia Elétrica ANO: 2025

Depto. de Engenharia Elétrica (ENE) - FT
Universidade de Brasília (UnB)
Campus Darcy Ribeiro
CEP 70919-970 - Brasília - DF - Brasil

DEDICATÓRIA

Dedico este trabalho aos meus pais, que me ensinaram o valor do esforço e da honestidade.

À minha noiva, aos meus filhos e ao meu neto, razão maior da minha perseverança e esperança no futuro.

AGRADECIMENTOS

Agradeço, com profundo respeito e admiração, ao meu orientador, Prof. Dr. João José Costa Gondim, por sua orientação firme, generosa e comprometida ao longo de todo este percurso. Sua confiança no potencial desta pesquisa e sua escuta atenta foram fundamentais para que este trabalho ganhasse forma e profundidade.

Ao meu coorientador, Prof. Dr. Robson de Oliveira Albuquerque, sou grato pela contribuição crítica, pelas sugestões precisas e pelo apoio constante, que enriqueceram significativamente esta dissertação.

A ambos, meu sincero reconhecimento por compartilharem conhecimento, paciência e tempo — elementos que fizeram toda a diferença nesta jornada.

A todos os meus familiares e amigos, por estarem presentes em cada conquista, mesmo nas horas silenciosas.

E, com profunda reverência:

A Exu, senhor das encruzilhadas e mensageiro entre os mundos, que permitiu que cada passo encontrasse passagem. *Laroyê!*

A Ogum, engenheiro cósmico, que abriu trilhas onde havia mata fechada e sustentou minha marcha com coragem e firmeza. *Ogunhê!*

E a Iansã, senhora do tempo e das transformações, que me ensinou a dançar com os ventos e atravessar as tempestades com dignidade, movimento e fé. *Eparrey oyá!*

RESUMO

O crescimento das ameaças cibernéticas tornou a Inteligência de Ameaças Cibernéticas (CTI) um elemento fundamental para a defesa proativa das organizações. A Inteligência de Fontes Abertas (OSINT) oferece um vasto volume de dados capazes de agregar contexto e profundidade às análises de CTI, mas o aproveitamento efetivo dessas informações ainda enfrenta desafios de filtragem, validação e aplicação estratégica. Este trabalho propõe e avalia um framework estruturado para o enriquecimento de CTI com dados provenientes de OSINT, fundamentado no modelo 5W3H. A metodologia desenvolvida organiza o processo de coleta, verificação e categorização dos dados, otimizando a identificação e a contextualização de Indicadores de Comprometimento (IoCs), como endereços IP associados a atividades maliciosas. Os experimentos realizados, envolvendo abordagens manuais, automatizadas e plataformas reconhecidas de threat intelligence, demonstraram que a aplicação sistemática do framework possibilita ganhos expressivos em completude, precisão e utilidade operacional dos relatórios de CTI. Conclui-se que a adoção de uma metodologia multidimensional e estruturada para o uso de OSINT no enriquecimento de CTI contribui de forma significativa para o fortalecimento da resiliência cibernética organizacional, apontando caminhos para melhores práticas e futuras evoluções na área.

ABSTRACT

The growth of cyber threats has made Cyber Threat Intelligence (CTI) a fundamental element for the proactive defense of organizations. Open Source Intelligence (OSINT) offers a vast volume of data capable of adding context and depth to CTI analyses; however, effectively leveraging this information still faces challenges related to filtering, validation, and strategic application. This work proposes and evaluates a structured framework for enriching CTI with OSINT data, based on the 5W3H model. The developed methodology organizes the process of data collection, verification, and categorization, optimizing the identification and contextualization of Indicators of Compromise (IoCs), such as IP addresses associated with malicious activities. The experiments conducted—covering manual, automated, and widely used threat intelligence platforms—demonstrated that the systematic application of the framework enables significant improvements in the completeness, accuracy, and operational utility of CTI reports. It is concluded that adopting a multidimensional and structured methodology for the use of OSINT in CTI enrichment significantly contributes to strengthening organizational cyber resilience, highlighting best practices and future directions in the field.

SUMÁRIO

LISTA DE FIGURAS

LISTA DE TABELAS

1 INTRODUÇÃO

A Inteligência de Ameaças Cibernéticas (Cyber Threat Intelligence – CTI) desempenha um papel crucial na defesa contra ataques cibernéticos, permitindo que as organizações compreendam e respondam proativamente a ameaças emergentes [3]. O cenário de ameaças cibernéticas intensificou-se: segundo o Verizon Data Breach Investigations Report 2024 (4), a exploração de vulnerabilidades como vetor de ataque inicial quase triplicou, com um aumento de 180% em relação ao ano anterior, enquanto ataques por meio da cadeia de suprimentos cresceram 68% (atingindo 15% das violações) (5). Adicionalmente, o IBM X-Force Threat Intelligence Index 2024 (6) destaca que o uso de credenciais válidas disparou 71% ano a ano, respondendo por 30% de todos os incidentes atendidos. Neste cenário, o desenvolvimento e a aplicação de estratégias de CTI tornaram-se indispensáveis não apenas para a mitigação de riscos, mas também para a antecipação de movimentos adversários, pois a crescente sofisticação e velocidade de evolução das ameaças exige que as organizações adotem processos de inteligência capazes de transformar dados dispersos em conhecimento acionável. A CTI permite compreender o contexto das ameaças, identificar padrões de comportamento de agentes hostis e antecipar suas prováveis ações, o que amplia a capacidade de resposta e reduz a superfície de ataque. Relatórios como o IBM X-Force Threat Intelligence Index 2024 e o Verizon DBIR 2024 apontam que adversários têm explorado cada vez mais vetores de ataque inovadores e coordenados, reforçando que estratégias baseadas em inteligência são fundamentais para orientar decisões de segurança, priorizar recursos e alinhar defesas às ameaças mais relevantes.

A crescente digitalização dos processos empresariais, impulsionada pela transformação digital e pela integração de tecnologias emergentes, tem ampliado significativamente a superfície de ataque das organizações. Segundo o International Data Corporation (7), os investimentos globais em transformação digital devem alcançar US$ 3,9 trilhões até 2027, refletindo a rápida incorporação de sistemas conectados e serviços baseados em nuvem em múltiplos setores. Essa interconectividade global, embora traga ganhos operacionais, também expõe organizações a um espectro mais amplo de ameaças — desde ataques de ransomware e botnets até campanhas avançadas e persistentes (APT) — conforme destacado pelo ENISA Threat Landscape 2023, que aponta a intensificação e diversificação desses vetores nos últimos anos.. Para enfrentar tal cenário, é vital que as empresas implementem estratégias de segurança baseadas em dados constantemente atualizados e analisados de modo crítico. No entanto, o Comitê Consultivo Nacional de Telecomunicações e Segurança (National Security Telecommunications Advisory Committee – NSTAC) destaca que, apesar da necessidade de colaboração, diversas equipes de cibersegurança do setor privado ainda hesitam em compartilhar informações de CTI e outros dados relevantes. Essa resistência está associada, entre outros fatores, a preocupações com a exposição de vulnerabilidades, questões regulatórias e riscos à imagem corporativa [8].

Neste contexto de crescente demanda por inteligência qualificada, a Inteligência de Fontes Abertas (Open Source Intelligence – OSINT) surge como um componente fundamental. Sua rápida expansão tem sido impulsionada por avanços tecnológicos e pela ampla disponibilidade de dados digitais, além do fortalecimento de comunidades colaborativas e de ferramentas especializadas. Atualmente, a OSINT se faz presente em múltiplos domínios – cibersegurança, aplicação da lei, operações militares, governança

e até mesmo investigação jornalística. Com o advento da Inteligência Artificial (IA) e do aprendizado de máquina, a capacidade de processamento e análise em tempo real atingiu novos patamares, possibilitando o monitoramento massivo de fluxos de informação, a detecção de padrões e a extração de insights relevantes para diversas finalidades. Casos recentes, como o uso de OSINT aprimorado por IA durante a guerra Russo-Ucraniana, exemplificam como dados civis provenientes de mídias sociais podem ser rapidamente convertidos em inteligência operacional, ampliando a consciência situacional e ao mesmo tempo suscitando questões éticas quanto ao envolvimento de civis [9].

Para o emprego eficaz de ferramentas de OSINT, é essencial compreender suas aplicações, limitações e pontos fortes específicos. De acordo com o Manual de Campanha do Exército Brasileiro (10), a OSINT é uma disciplina de inteligência fundamentada na coleta, exploração e disseminação de dados públicos disponíveis (DPD) para atender a necessidades de inteligência, contribuindo para a consciência situacional e subsidiando outras disciplinas de obtenção. Van Puyvelde & Rienzi (11) a definem como um conjunto de práticas que envolve a coleta, validação e exploração de informações publicamente acessíveis para atender a requisitos específicos, ressaltando que a consolidação dessa disciplina depende de um corpo estruturado de conhecimento e de padrões reconhecidos.

Embora não haja consenso absoluto sobre uma tipologia única, Yadav et al. (12) apontam que, conforme a aplicação e a natureza da fonte, a OSINT pode abranger categorias como GEOINT (Inteligência Geoespacial), SOCMINT (Inteligência de Mídias Sociais), HUMINT (Inteligência Humana obtida de forma aberta), entre outras. Cada domínio apresenta potenciais e desafios próprios, como se vê na tabela 1.1. O êxito nas operações de OSINT depende da seleção criteriosa de fontes e ferramentas, da validação cruzada de dados e da observância de padrões legais e éticos durante todo o ciclo de coleta, análise e disseminação da informação.

Tabela 1.1: Categorias de OSINT, exemplos, aplicações e limitações.

Categoria	Definição / Escopo	Fontes	Aplicações	Limitações
GEOINT	Análise de dados geoespaciais e de localização.	Imagens de satélite, sensores remotos, mapas digitais.	Monitoramento ambiental, análise de ameaças, rastreamento de deslocamentos.	Baixa resolução, necessidade de validação, restrições legais.
SOCMINT	Dados de mídias sociais para padrões e interações.	Twitter/X, Facebook, Instagram, Telegram.	Rastreamento de desinformação, análise de opinião, redes de influência.	Alto volume, manipulação de dados, barreiras de privacidade.
HUMINT (aberta)	Informações obtidas de pessoas/eventos públicos.	Entrevistas, conferências, declarações oficiais.	Identificar relacionamentos e capacidades.	Viés, confiabilidade, risco de desinformação.
OSINF	Dados públicos brutos antes de análise OSINT.	Registros públicos, bases comerciais, sites governamentais.	Base documental para investigações.	Necessidade de validação e atualização.
IMINT (aberta)	Inteligência por imagens e vídeos públicos.	Fotos, câmeras públicas, YouTube.	Geolocalização, verificação de incidentes, monitoramento.	Manipulação digital, necessidade de verificação especializada.
SIGINT (aberta)	Sinais de comunicação abertos.	Rádio, TV, dados ADS-B.	Monitoramento de comunicações e tráfego.	Limites técnicos e legais.

O enriquecimento de CTI com dados de OSINT representa, assim, um salto qualitativo na produção de inteligência, ampliando o valor dos indicadores ao adicionar contexto e profundidade [13, 14]. Contudo, tal enriquecimento deve ser guiado por processos bem definidos, capazes de evitar a coleta de grandes volumes de dados irrelevantes ou não verificados, o chamado "lixo informacional" [9], entendido como o conjunto de registros que não agregam valor analítico ou induzem a conclusões equivocadas. Exemplos típicos incluem listas de IPs copiadas de repositórios públicos contendo entradas duplicadas, endereços privados (RFC1918) ou já desativados; hashes de malware de uso único, desatualizados ou sem vínculo claro com campanha/TTP; capturas massivas de postagens em redes sociais impregnadas de boatos, contas automatizadas e conteúdo fora de escopo; indicadores extraídos de pastebins sem fonte verificável ou com carimbo de data antigo; resultados de varreduras (p.ex., Shodan/Censys) que refletem mudanças temporárias de configuração, sem evidência de comportamento malicioso; e domínios "typosquatted" sem prova de uso operacional por atores de ameaça. Para mitigar esse ruído, o processo deve incluir normalização e deduplicação, checagens de atualidade e verificação cruzada de fontes, além do registro de proveniência e de critérios explícitos de aceitação/rejeição de indicadores [15]. Para tanto, destaca-se a necessidade de validação criteriosa de dados e fontes [16, 1], bem como de filtragem sistemática e estruturação das informações de maneira a descrever de forma precisa não apenas a ameaça, mas também os principais

atores e circunstâncias que a tornam viável [17]. Um dos métodos que se mostram eficazes para essa finalidade é o "5W3H", que proporciona um modelo de organização baseado em oito dimensões essenciais para compreensão e comunicação de eventos complexos [18].

O presente estudo foi concebido para abordar uma lacuna crítica da cibersegurança contemporânea: a eficácia do uso de OSINT para o enriquecimento de informações de CTI. A premissa central é que, enquanto o volume de dados cresce exponencialmente, a capacidade de filtrá-los, verificá-los e aplicá-los de modo estratégico não tem acompanhado esse ritmo (19). A adoção de uma metodologia robusta para integrar OSINT ao processo de CTI tende a elevar a qualidade, a precisão e a utilidade dos produtos de inteligência, pois alinha a coleta e o enriquecimento de indicadores a práticas reconhecidas de gestão de risco e decisões informadas. O NIST CSF 2.0 orienta que organizações integrem fontes de inteligência de ameaças ao ciclo de gestão (funções Govern/Identify/Detect/Respond), como forma de priorizar controles e reduzir exposição, reforçando a necessidade de processos formais para transformar dados em ação, um pilar de resiliência cibernética (20). Em paralelo, a ISO/IEC 27001:2022 (A.5.7 – Threat intelligence) torna explícita a expectativa de coletar, analisar e produzir inteligência sobre ameaças relevantes ao negócio, conectando esse insumo a medidas preventivas, detectivas e corretivas — isto é, resiliência operacional diante de cenários mutáveis (21).

Evidências empíricas e revisões apontam que a CTI melhora a situational awareness, auxilia a priorização (reduzindo ruído e falsos positivos) e acelera detecção/mitigação, especialmente quando a informação é compartilhada e contextualizada entre múltiplas fontes, exatamente o papel de uma integração OSINT bem estruturada (22). Relatórios setoriais como o ENISA Threat Landscape reforçam que, diante do aumento e da sofisticação das ameaças, resiliência depende de inteligência acionável e tempestiva, capaz de orientar resposta e continuidade do negócio (23). Em suma, quando OSINT é incorporado por meio de processos definidos (requisitos de inteligência, validação cruzada, contextualização e disseminação dirigida), as organizações priorizam melhor seus esforços, reduzem a incerteza operacional e recuperam-se mais rapidamente de incidentes — atributos centrais de resiliência.

Hipótese (H1): a adoção do framework 5W3H integrado a OSINT aumenta a qualidade, a precisão e a utilidade operacional dos produtos de CTI, em comparação a abordagens manuais e a plataformas OSINT generalistas. A eficiência será avaliada em estudo de caso com dois IPs analisados em quatro cenários (manual, EnricherV2, OTX AlienVault e Pulsedive), pelas métricas: (i) cobertura 5W3H; (ii) profundidade/qualidade do conteúdo; (iii) usabilidade/complexidade operacional; e (iv) observações sobre falsos positivos/negativos e reprodutibilidade.

1.1 OBJETIVO GERAL

Desenvolvimento de um framework para o enriquecimento de Inteligência de Ameaças Cibernéticas (CTI) a partir da integração sistemática de dados provenientes de Inteligência de Fontes Abertas (OSINT) visando aumentar a precisão, a relevância e a utilidade das informações de CTI no apoio à tomada de decisão em segurança cibernética.

1.1.1 Objetivos Específicos

1. Propor uma metodologia estruturada para organizar e direcionar o enriquecimento de informações de CTI com dados obtidos por meio de OSINT.

2. Desenvolver uma ferramenta computacional capaz de coletar, correlacionar e estruturar automaticamente dados de fontes abertas, integrando-os aos parâmetros definidos pelo framework proposto.

3. Avaliar a eficácia da metodologia e da ferramenta, por meio de experimentos, demonstrando ganhos em contextualização, precisão e acionabilidade dos produtos de inteligência.

4. Identificar desafios, limitações e melhores práticas na integração de dados OSINT ao processo de enriquecimento de CTI, oferecendo subsídios para a adoção da abordagem proposta por organizações que atuam em defesa cibernética.

1.2 CONTRIBUIÇÕES DESTE TRABALHO

A primeira contribuição deste trabalho consiste na apresentação de uma revisão bibliográfica abrangente, que mapeou as principais abordagens, desafios e oportunidades no campo da inteligência de ameaças cibernéticas (CTI), enriqeucimento de dados e inteligência de fontes abertas (OSINT). Essa revisão fundamenta o estado da arte e serve como referência para pesquisadores e profissionais interessados em metodologias de integração e automação de processos de CTI, enriquecimento e OSINT.

Em seguida, o trabalho propôs uma metodologia estruturada para o enriquecimento de CTI. Esta proposta permitiu organizar e priorizar o processo de coleta, correlação e análise de dados provenientes de OSINT, resultando em maior precisão, relevância e aplicabilidade dos produtos de inteligência em ambientes operacionais de cibersegurança. Diferentemente de abordagens já existentes, o framework aqui desenvolvido combina um modelo 5W3H com integração automatizada de múltiplas fontes OSINT, permitindo enriquecimento contextual e geração de relatórios com atributos completos para tomada de decisão. Além disso, sua arquitetura foi pensada para ser adaptável a diferentes plataformas e cenários, incluindo contextos de alta criticidade e restrições de tempo, o que amplia sua aplicabilidade prática e o diferencia de soluções que focam apenas em coleta ou análise isolada.

Como desdobramento prático, parte dos resultados foi consolidada no artigo (24) apresentado no WorldCIST 2025, no dia 15 de abril de 2025. Ressalta-se, no entanto, que até a data de conclusão desta dissertação, o artigo ainda não havia sido oficialmente publicado nos proceedings do evento. Outro artigo, submetido e aceito para publicação no CISTI 2025, também resultou da pesquisa realizada, mas não pôde ser apresentado nem publicado em razão de limitações de financiamento.

Por fim, destaca-se o desenvolvimento e o registro de uma ferramenta computacional (25) que operacionaliza a metodologia proposta, automatizando o enriquecimento de dados de diversas fontes OSINT conforme os parâmetros da metodologia proposta. O registro do software garante sua proteção intelectual e abre caminho para sua utilização controlada em contextos acadêmicos e profissionais, ampliando o impacto e a aplicabilidade das contribuições desta pesquisa.

1.3 ESTRUTURA DO TRABALHO

Este trabalho está organizado da seguinte forma: O Capítulo 2 – FUNDAMENTAÇÃO TEÓRICA E TRABALHOS CORRELATOS: Revisita os principais conceitos de CTI, OSINT, plataformas de inteligência de ameaças (TIP), ciclo de inteligência, indicadores de comprometimento (IoCs), além de discutir trabalhos correlatos e os desafios atuais do tema; 0 Capítulo 3 – DISCUSSÃO DO PROBLEMA E PROPOSTA: Discute o problema a ser resolvido, apresenta a metodologia e detalha o framework desenvolvido, apresentando a modelagem dos parâmetros, a integração das fontes de dados, as características da ferramenta construída e a comparação com métodos existentes; O Capítulo 4 – RECURSOS E MÉTODOS apresenta os principais recursos e métodos usados na implementação do framework, incluindo ferramentas, plataformas e feeds empregados no enriquecimento da inteligência de ameaças cibernéticas. Também compara a ferramenta criada com abordagens tradicionais e recentes de CTI, destacando diferenças, vantagens e limitações em relação aos métodos existentes. ; No Capítulo 5 – RESULTADOS EXPERIMENTAIS temos a descrição dos experimentos realizados, dos cenários de aplicação, dos parâmetros utilizados, os resultados obtidos com o enriquecimento de IoCs e a discussão crítica sobre a eficácia do método proposto; Por fim, no Capítulo 6 – CONCLUSÃO apresentamos as principais conclusões, as limitações encontradas, sugestões de trabalhos futuros e recomendações para a adoção da metodologia.

2 FUNDAMENTAÇÃO TEÓRICA E TRABALHOS CORRELATOS

Este capítulo apresenta os principais conceitos, métodos e referências que fundamentam a proposta deste trabalho. São discutidos a evolução da inteligência de ameaças cibernéticas (CTI), os fundamentos e aplicações de inteligência de fontes abertas (OSINT), bem como as principais plataformas, processos e desafios envolvidos no enriquecimento de dados em CTI. Ao final, são revisados trabalhos correlatos e apontadas lacunas que motivam a abordagem adotada nesta pesquisa.

2.1 DEFINIÇÕES

2.1.1 Inteligência

A noção de inteligência, em seu sentido amplo, pode ser compreendida como o processo sistemático de coleta, avaliação, análise e disseminação de informações relevantes para apoiar processos decisórios, seja no contexto estatal, empresarial ou militar (26). Internacionalmente, autores como Sherman Kent, um dos precursores do pensamento sobre inteligência nos Estados Unidos, definem inteligência como o conhecimento vital para a tomada de decisão em situações de incerteza, obtido a partir da transformação de dados brutos em informação útil e acionável (27). Essa perspectiva é compartilhada por Michael Warner, que descreve inteligência como "informação coletada, analisada e distribuída de maneira oportuna, a fim de apoiar o tomador de decisão" (26).

No contexto brasileiro, a Doutrina da Atividade de Inteligência da Agência Brasileira de Inteligência (ABIN) destaca que a atividade de inteligência visa à produção de conhecimentos e à realização de ações voltadas à redução de vulnerabilidades e neutralização de ameaças contra a segurança das pessoas e instituições, bem como à proteção de informações sensíveis e à identificação de oportunidades de interesse público (28). Segundo a doutrina nacional, a atividade de inteligência estrutura-se em dois ramos — inteligência e contrainteligência — e dois elementos constitutivos: análise (produção de conhecimento) e operações (ações especializadas), sendo ambos fundamentais para a eficácia e a legitimidade do processo. A Doutrina (28) define o ramo Inteligência como o responsável por produzir e difundir conhecimentos estratégicos capazes de apoiar o processo decisório e antecipar ameaças ou oportunidades relevantes ao Estado e à sociedade. Já o ramo Contrainteligência é voltado à salvaguarda desses conhecimentos, bem como à proteção de pessoas, instalações e meios sensíveis, por meio da prevenção, detecção, obstrução e neutralização de ações adversas, como espionagem e sabotagem, que possam comprometer a segurança nacional.

Do ponto de vista metodológico, a inteligência diferencia-se de outras disciplinas informacionais por buscar não apenas descrever eventos, mas interpretá-los em seu contexto, antecipando cenários e subsidiando ações estratégicas. Como exemplificado em (1), o conceito de inteligência envolve múltiplas fases:

coleta, processamento, análise e difusão de informações, com ênfase na ética, no rigor metodológico e na responsabilidade social. A literatura contemporânea reconhece que a obtenção e a análise de dados — que podem ser provenientes de fontes abertas, técnicas, humanas ou técnicas especializadas — só se convertem em inteligência quando submetidos a processos de validação, integração e contextualização, promovendo a redução da incerteza em ambientes complexos.

No campo da segurança cibernética, a definição de inteligência mantém-se alinhada aos princípios tradicionais, porém incorpora características operacionais específicas que refletem as necessidades práticas desse domínio. Essas características referem-se à aplicação direta do conhecimento produzido para orientar ações concretas, como a detecção, mitigação e resposta a incidentes, dentro de janelas de tempo muito reduzidas. Em um ambiente marcado por grandes volumes de dados e pela rápida evolução de ameaças e oportunidades (29), a inteligência cibernética precisa ser processada, contextualizada e transformada em recomendações acionáveis com agilidade. Dessa forma, ela se torna um elemento essencial para antecipar riscos e sustentar respostas fundamentadas, representando a adaptação dos conceitos clássicos às demandas dinâmicas e complexas do cenário digital contemporâneo.

2.1.2 Ciclos de Inteligência

O ciclo de inteligência é tradicionalmente compreendido como um processo dinâmico e iterativo, estruturado para transformar dados brutos em conhecimento útil para a tomada de decisão. O ciclo clássico — adotado por diversas agências e amplamente citado na literatura — compreende as fases de Direção e Planejamento, Coleta, Processamento, Análise e Difusão, e sua lógica fundamenta grande parte das doutrinas nacionais e internacionais (27, 26).

Figura 2.1: As etapas do Ciclo de Inteligência genérico. Fonte: (1)

Além do modelo clássico, outros ciclos de inteligência são reconhecidos mundialmente. O UK Intelligence Cycle, utilizado pelas principais agências britânicas, enfatiza quatro grandes etapas: Direção,

Coleta, Processamento e Disseminação, destacando o fluxo de feedback contínuo entre cada fase e a necessidade de constante reavaliação das prioridades (30). Já o modelo do Intelligence Community dos Estados Unidos, formalizado pelo Office of the Director of National Intelligence (ODNI), estrutura o processo em cinco fases, constantemente utilizando uma sexta: Planejamento e Direção, Coleta, Processamento e Exploração, Análise e Produção, Disseminação e Integração, Avaliação e Feedback, reforçando a integração dos resultados e o papel da avaliação para a melhoria contínua. Tal ciclo pode ser visto , por exemplo, no Space Doctrine Publication 2-0, Intelligence (31).

Tanabe et al. (1) mostram que, mesmo quando se foca apenas em OSINT, o processo não se esgota na coleta: técnicas de verificação, conversão de formatos e análise precisam ser integradas ao ciclo para que a informação ganhe valor operacional. Já Silva et al. (32) reforçam que a mesma lógica de etapas sucessivas é indispensável à produção de CTI de qualidade; ao acoplar métricas de enriquecimento ao modelo clássico, os autores evidenciam que a fase de planejamento orienta todo o restante do fluxo e impacta diretamente na completude e na precisão do produto final.

Pesquisas recentes têm ampliado o debate ao aproximar diferentes disciplinas de coleta. Macêdo et al. (33) destacam que a convergência entre HUMINT e OSINT adiciona camadas contextuais cruciais, sobretudo nos estágios de verificação e análise do ciclo, pois permite qualificar a confiabilidade das fontes públicas ao cruzá-las com evidências humanas. Essa combinação robustece a etapa de avaliação e realimenta o planejamento com insumos mais ricos, fechando o “feedback loop” de maneira mais eficiente.

Além das adaptações disciplinares, alguns trabalhos propõem metodologias específicas para melhorar cada fase do ciclo. Pincovscy e Gondim (29) demonstram que, ao inserir sensores de rede na etapa de coleta e acoplar um processo automatizado de enriquecimento antes mesmo do armazenamento em plataformas de compartilhamento (TISP), obtém-se ganho expressivo de velocidade e relevância na análise subsequente. Esses resultados reforçam a premissa de que o ciclo de inteligência não é um modelo estático: ele pode (e deve) ser customizado de acordo com o domínio de aplicação, contanto que se mantenha a lógica iterativa planejamento-coleta-análise-disseminação-avaliação como eixo estruturante.

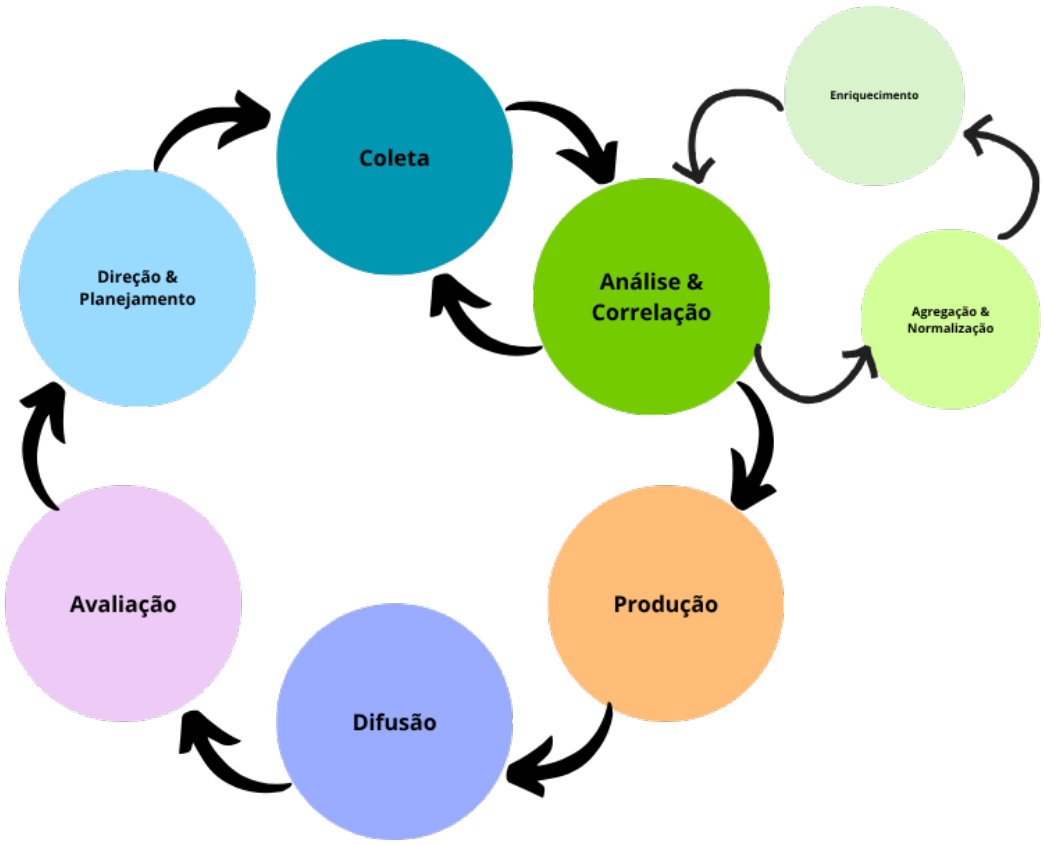

Figura 2.2: As etapas do Ciclo de Inteligência para a presente Metodologia. Fonte: Proposta do Autor

Para o presente trabalho, o modelo a ser considerado de ciclo de inteligência é o proposto na Figura 2.2. O fluxo inicia em Direção & Planejamento, onde se definem os requisitos de inteligência e as fontes a serem consultadas. A Coleta reúne dados brutos que, em seguida, entram num micro-loop iterativo: primeiro passam por Agregação & Normalização, onde registros duplicados são consolidados; depois por Enriquecimento, que acrescenta contexto externo (reputação, táticas, geolocalização); e finalmente por Análise & Correlação, onde padrões e relacionamentos são extraídos. Caso a análise revele lacunas o processo retrocede dentro desse loop até que a informação esteja completa. Com os achados consolidados, a fase de Produção gera relatórios e artefatos táticos, que seguem para Difusão e implantação em equipes de defesa. Por fim, a Avaliação mede a utilidade do produto (ex.: redução de falsos-positivos) e realimenta o planejamento, fechando o ciclo e impulsionando melhorias contínuas na geração de CTI.

2.1.3 Inteligência de Fontes Abertas (Open Source Intelligence - OSINT)

A Inteligência de Fontes Abertas (OSINT) consiste na coleta sistemática, processamento crítico e análise de informações publicamente disponíveis, incluindo websites, mídias sociais, repositórios governamentais, arquivos multimídia e outros meios abertos, com o propósito de sustentar decisões estratégicas e operacionais. No contexto militar brasileiro, o Manual de Campanha MC 2.40-54 (10) define OSINT como a disciplina de inteligência baseada em dados coletados de fontes de caráter público, tais como meios de comunicação tradicionais e conteúdos disponíveis no espaço cibernético, ressaltando seu papel fundamental na produção de conhecimentos e na construção de consciência situacional.

Estudos estimam que entre 80% e 90% do volume total de inteligência consumido por agências governamentais e corporações já deriva de fontes abertas, fenômeno impulsionado pela explosão de dados digitais e pela maturidade de técnicas de data mining e aprendizado de máquina (9). O avanço da internet ampliou tanto o volume quanto a diversidade das fontes, estendendo o alcance da OSINT para domínios como defesa, segurança cibernética, resposta a desastres, investigação de crimes e análise de ameaças emergentes (34).

O ciclo de inteligência tradicional — direção, coleta, processamento, análise, disseminação e retroalimentação — posiciona a OSINT como etapa essencial de coleta, embora seus insumos permeiem todas as fases do processo quando métodos adequados são empregados. O manual do Exército Brasileiro reforça que a exploração de fontes abertas frequentemente precede outras disciplinas de inteligência, subsidiando o planejamento e a execução das demais atividades (10). Tanabe et al. (1) propõem um workflow integrado em que técnicas de busca, extração, preservação e verificação são mapeadas diretamente sobre cada etapa do ciclo, garantindo rastreabilidade e confiabilidade desde a definição das perguntas de inteligência até a entrega do produto final.

A literatura internacional mostra que a OSINT evoluiu para além da coleta passiva de dados, incorporando técnicas avançadas de mineração de texto, aprendizado de máquina, análise semântica e inteligência artificial para a extração e enriquecimento de informações relevantes (34). Chaudhary e Bansal (2022) apresentam um ciclo de extração de OSINT dividido em três fases principais: aquisição de dados, enriquecimento de dados e inferência de conhecimento, ressaltando os desafios impostos pela heterogeneidade, volume e velocidade dos dados disponíveis.

Ferramentas OSINT tendem a se tornarem obsoletas rapidamente devido a mudanças em APIs, políticas de privacidade ou modelos de negócios. Por isso, a perenidade do OSINT reside nas técnicas, não nos aplicativos: web scraping orientado a padrões, resolução de metadados, análise geoespacial, correlação temporal, enriquecimento semântico e vetorização de texto são habilidades que permanecem válidas independentemente da plataforma utilizada (1).

O emprego de OSINT apresenta benefícios e desafios específicos. Por um lado, facilita o acesso a grandes volumes de dados com custos e riscos reduzidos; por outro, impõe a necessidade de avaliações rigorosas de credibilidade, relevância e atualização das fontes, além de implicações ético-legais relevantes (10, 34). Além disso, a integração de OSINT com outras disciplinas, como HUMINT, tem se mostrado estratégica para qualificar a confiabilidade das informações e ampliar a cobertura analítica (33).

Em síntese, OSINT consolida-se como componente indispensável para programas modernos de inteligência, agregando valor à segurança nacional, à ciberdefesa, à análise de ameaças e à resiliência organizacional. Sua efetividade, contudo, depende de processos bem estruturados, adoção de técnicas robustas e arcabouço normativo que assegure a legalidade e a ética das operações.

2.1.4 Inteligência de Ameaças

A inteligência de ameaças (Threat Intelligence) é tradicionalmente compreendida como o processo sistemático de coleta, análise e interpretação de informações sobre atores hostis, eventos, capacidades e potenciais riscos que possam afetar a segurança de pessoas, ativos ou organizações. O objetivo fundamental

dessa disciplina é transformar dados dispersos e, muitas vezes, fragmentados em conhecimento contextualizado e acionável para subsidiar decisões táticas, operacionais e estratégicas, promovendo antecipação e resiliência diante de cenários incertos (35).

No contexto contemporâneo, a inteligência de ameaças expande seu escopo para além do monitoramento de ações imediatas de adversários, buscando identificar tendências, padrões de comportamento, relações de causa e efeito, e fatores geopolíticos ou sociais que possam influenciar o surgimento de novas ameaças. Essa perspectiva orienta tanto ações preventivas quanto respostas ágeis a eventos já em andamento, sendo reconhecida como elemento central para a gestão integrada de riscos.

A produção efetiva de inteligência de ameaças envolve múltiplos níveis de análise — da coleta bruta de dados ao refinamento e validação de hipóteses — e pode fazer uso de diferentes fontes, incluindo fontes abertas, relatos de campo, documentos públicos, relatórios de incidentes, ou qualquer insumo relevante à caracterização do contexto de ameaça. O ciclo dessa inteligência é tipicamente iterativo: definição de requisitos, coleta, avaliação, análise, síntese e disseminação, com feedback contínuo para aprimoramento das práticas e atualização dos produtos de inteligência.

O valor da inteligência de ameaças reside na sua capacidade de antecipar riscos emergentes, orientar a alocação eficiente de recursos defensivos e apoiar a tomada de decisões embasadas em conhecimento, contribuindo para a resiliência organizacional e para a redução da incerteza em ambientes cada vez mais complexos e interconectados.

2.1.5 Inteligência de Ameaças Cibernéticas

A inteligência contra ameaças cibernéticas (Cyber Threat Intelligence – CTI) é definida como um processo contínuo de coleta, análise e disseminação de informações relevantes sobre ameaças conhecidas ou emergentes, com o propósito de fortalecer a segurança de sistemas e apoiar decisões estratégicas e operacionais no enfrentamento de riscos digitais. De acordo com Osliak et al. (13), a CTI constitui uma ferramenta fundamental para proteger infraestruturas críticas, permitindo a antecipação a ataques por meio do monitoramento constante de indicadores e comportamentos suspeitos. Esses dados incluem informações contextuais sobre os agentes maliciosos, suas motivações, técnicas e objetivos, além de evidências técnicas como endereços IP, hashes de arquivos maliciosos e domínios comprometidos.

Sun et al. (15) complementam essa visão ao destacar que a CTI é composta não apenas por informações observáveis, como os Indicadores de Comprometimento (IoCs), mas também por componentes analíticos que fornecem contexto e implicações. Esses dados são organizados e representados por meio de padrões como STIX e TAXII, que facilitam o intercâmbio estruturado de informações entre diferentes sistemas e organizações, promovendo uma visão integrada e confiável do cenário de ameaças cibernéticas.

Um dos principais elementos que ampliam a utilidade da CTI é o processo de enriquecimento dos IoCs (enrichment), que consiste em associar aos indicadores informações adicionais relevantes ao seu contexto de ocorrência. Isso pode incluir dados como geolocalização do endereço IP malicioso, reputação do domínio, CVEs (Common Vulnerabilities and Exposures) exploradas, vínculos com campanhas anteriores e padrões de comportamento do atacante. Conforme Osliak et al. (13), o enriquecimento desempenha papel decisivo na redução de falsos positivos, na melhora da acurácia na priorização de alertas e na expli-

cabilidade das decisões de resposta e mitigação, conferindo rastreabilidade e justificativa técnica às ações tomadas.

Sun et al. (15) destacam que, especialmente em ambientes dinâmicos e de grande escala, como redes corporativas e infraestruturas industriais, o enriquecimento automatizado de IoCs é essencial para viabilizar uma resposta rápida e eficaz. A agregação de dados contextuais permite compreender não apenas o que aconteceu, mas também como e por que aconteceu, fornecendo uma base robusta para ações defensivas orientadas por inteligência. Assim, o valor da CTI não reside apenas na detecção de ameaças, mas na sua capacidade de apoiar decisões informadas com base em dados enriquecidos e validados.

Em síntese, a CTI fornece uma estrutura crítica para fortalecer a resiliência cibernética das organizações, permitindo uma compreensão profunda das ameaças e facilitando a construção de defesas adaptativas, especialmente quando se integra ao ciclo de vida da resposta a incidentes e ao gerenciamento contínuo de riscos cibernéticos.

2.1.6 Plataforma de Inteligência de Ameaças (Threat Intelligence Platform - TIP)

Uma Plataforma de Inteligência de Ameaças (Threat Intelligence Platform – TIP) é, em essência, um sistema sociotécnico automatizado cujo propósito é transformar dados heterogêneos sobre ameaças em inteligência acionável. A literatura especializada converge em cinco funções mínimas que caracterizam esse tipo de plataforma:

1. Coleta e ingestão multiorigem – feeds OSINT, relatórios técnicos, sensores internos e APIs externas são consumidos de maneira contínua; essa automação é apontada como requisito para defesa em tempo real.(36)
2. Normalização e padronização – indicadores brutos (IoCs) são convertidos para modelos comuns (p.ex., STIX/TAXII), superando a heterogeneidade de formatos observada nas fontes de CTI.(37)
3. Enriquecimento e correlação contextual – dados coletados são combinados com telemetria da própria infraestrutura para gerar IoCs compostos e calcular scores de risco, elevando sua relevância operacional.(16)
4. Análise e priorização – técnicas de mineração e aprendizado de máquina classificam, ranqueiam e identificam lacunas ou sobreposição entre fontes, possibilitando foco nas ameaças mais críticas.(38)
5. Disseminação e integração – a inteligência resultante é distribuída a SIEMs, IDS/IPS, equipes SOC/-CERT e parceiros externos via APIs ou canais de compartilhamento confiáveis, fechando o ciclo de ação.

Assim, para fins desta dissertação, define-se TIP como um ambiente integrado de processos, modelos de dados e serviços que viabiliza:

- (i) ingestão automatizada,
- (ii) harmonização semântica,

- (iii) correlação contextual,
- (iv) priorização orientada a risco e
- (v) compartilhamento seguro de inteligência, sustentando respostas proativas a ameaças cibernéticas.

2.1.7 Observáveis e Indicadores de Comprometimento (IoCs)

Observáveis são evidências discretas em um ambiente digital — endereços IP, hashes de arquivos, nomes de domínio, chaves de registro, strings de user-agent, entre outros — que, por si sós, apenas sugerem uma anomalia. Quando vinculados a uma ameaça confirmada, esses artefatos tornam-se Indicadores de Comprometimento (IoCs): sinais de alto grau de confiança de que ocorreu (ou está em curso) uma violação de segurança. No estágio inicial da defesa cibernética, IoCs serviram sobretudo a respostas automatizadas, bloqueando IPs maliciosos ou isolando hosts infectados.

Contudo, ataques contemporâneos, em especial aqueles conduzidos por Ameaças Persistentes Avançadas (APT), já não se deixam capturar por listas estáticas de IoCs. Hagen e Helkala (39) mostram que a sofisticação atual inclui mecanismos furtivos de persistência, exfiltração gradual de dados e táticas de evasão que exigem conjuntos dinâmicos de indicadores, frequentemente derivados de padrões de comportamento e não apenas de artefatos individuais. O estudo revisa 30 anos de incidentes e evidencia a transição de IoCs simples (IP ou hash único) para "IoCs compostos" que combinam temporalidade, sequência de eventos e correlação com táticas, técnicas e procedimentos (TTPs).

Nesse cenário, a comunidade científica propõe abandonar a visão atomizada dos IoCs e adotar um enfoque sistêmico. O artigo (40) descreve um método baseado em teoria dos grafos no qual cada artefato é modelado como vértice e as relações entre eles são arestas ponderadas; o resultado é um "IoC de sistema" capaz de representar toda a cadeia de intrusão, permitindo comparar incidentes via (parcial) isomorfismo de grafos. Essa representação não só integra múltiplos observáveis (rede, host, aplicação) como quantifica peso e direção das interações, fornecendo inteligência mais rica para priorização de respostas.

Portanto, nesta dissertação, IoCs serão tratados em dois níveis complementares:

1. Nível artefato — observáveis isolados, úteis para bloqueios rápidos e detecção assinada;
2. Nível sistêmico — modelos gráficos que capturam relações, temporalidade e contexto operacional, alimentando análises manuais, aprendizado de máquina e modelos preditivos.

Essa abordagem reconhece que a eficácia da Inteligência de Ameaças depende não apenas de detectar "o quê" (hash, IP), mas de compreender "como" e "por que" os artefatos se combinam em campanhas sofisticadas, alinhando-se às recomendações recentes da literatura sobre a complexidade dos IoCs e a necessidade de métodos analíticos avançados.

2.1.8 Indicadores de Ataque (IoAs)

Indicadores de Ataque (IoAs) constituem a evolução natural dos tradicionais IoCs. Enquanto IoCs descrevem "o que aconteceu" (ex.: hash de malware, endereço IP malicioso), os IoAs descrevem "o que

está acontecendo" — isto é, padrões comportamentais, sequências de eventos e contexto operacional que revelam a progressão tática de um adversário antes que o impacto se consolide.

1. Base nos TTPs do adversário: A taxonomia MITRE ATT&CK fornece o vocabulário operacional para capturar táticas, técnicas e procedimentos; cada IoA mapeia uma ou mais fases da cadeia de ataque (por exemplo "Privileged Escalation → Credential Dumping → Lateral Movement"), permitindo inferir intenções futuras em vez de apenas registrar artefatos legados.

2. Correlação temporal e contextual. IoAs combinam telemetria em tempo real — como logs de sistema, fluxo de rede e eventos de autenticação — para estabelecer relações de sequência, duração e coocorrência entre ações suspeitas. Ao observar a ordem dos eventos e o contexto em que ocorrem (horário, origem, privilégio envolvido, dependências de serviço, etc.), é possível reconhecer padrões característicos de ataques em andamento ou prestes a acontecer, mesmo quando não há um IoC previamente conhecido.

3. Resposta proativa e orquestrada. Ao detectar uma cadeia de TTPs em formação, o SOC pode interromper o "kill chain" —bloqueando credenciais recém-criadas, isolando hosts em quarentena ou ativando controle de acesso adaptativo— antes que o atacante complete suas metas de persistência ou exfiltração. Estudos que analisam três décadas de APTs comprovam que essa visão dinâmica é crucial para lidar com persistência avançada, movimentos laterais e técnicas de evasão que já não se manifestam em IoCs simples (39).

4. Integração com modelos analíticos. Uma vez organizados em bases de dados históricas, os IoAs alimentam algoritmos de aprendizado de máquina, correlações estatísticas e regras heurísticas capazes de prever a progressão do ataque, atribuir níveis de criticidade e sugerir respostas priorizadas. A inclusão de variáveis como relevância do ativo, frequência de ocorrência e impacto potencial aumenta a precisão dos alertas e reduz falsos-positivos, oferecendo às equipes de segurança tempo hábil para ações defensivas.

Em resumo, um IoA é um conjunto organizado de relações temporais e contextuais entre TTPs que, quando observadas em tempo de execução, sinalizam um ataque em curso ou iminente. Diferentemente dos IoCs, os IoAs descrevem comportamentos e não apenas artefatos, capacitando detecções preditivas e defesas acionáveis antes da materialização do impacto.

2.1.9 Enriquecimento de Dados

O enriquecimento de dados (data enrichment) é a etapa que converte artefatos técnicos isolados (por exemplo, um endereço IP suspeito ou o hash de um executável) em inteligência contextualizada capaz de orientar a tomada de decisão. Esse processo é especialmente crítico em um cenário marcado pela alta elasticidade das redes, nas quais endereços, serviços e configurações mudam rapidamente, e pelo uso único ou de curta duração de malwares. Nessas condições, indicadores técnicos brutos tornam-se rapidamente obsoletos, exigindo que sejam complementados com informações adicionais, como origem, comportamento, relações com outras ameaças e associação a campanhas específicas. Assim, o enriquecimento aumenta

a relevância e a utilidade operacional dos dados, permitindo priorizar respostas, correlacionar eventos e adotar medidas mais assertivas de mitigação. Esse processo ocorre em três frentes complementares:

1. Contexto socio-organizacional. Acrescentar metadados sobre quem controla um ativo amplia drasticamente o valor analítico do indicador. Moriot et al. (41) demonstram um algoritmo que cruza RDAP com Wikidata para rotular endereços IP segundo tipo de entidade, ramo de atividade e porte (número de empregados). O resultado é uma caracterização que distingue, por exemplo, tráfego proveniente de provedores residenciais, universidades ou grandes nuvens públicas, reduzindo falsos-positivos e permitindo contramedidas sob medida.

2. Harmonização semântica e fusão multiorigem. Em ambientes corporativos, IoCs chegam por canais diversos (sensores internos, feeds OSINT, bases governamentais e redes sociais). Pesquisas recentes (42) demonstram que um pipeline modular com etapas de descoberta, normalização, curadoria e vinculação pode transformar esses fluxos heterogêneos em um repositório semântico unificado. Ao acrescentar anotações de contexto e regras de inferência, o processo incorpora metadados de proveniência, qualidade e temporalidade, ampliando a consciência situacional e viabilizando consultas complexas em tempo quase real, mesmo diante de grandes volumes de dados.

3. Abstração progressiva e aprendizagem automática. Para que os dados enriquecidos sejam realmente acionáveis, eles precisam evoluir de leituras brutas para representações compreensíveis por humanos e máquinas. Park et al. (43) propõem um esquema extensível que descreve perfis, procedimentos de abstração e funções de controle em pilha semântica; sensores IoT, por exemplo, têm seus valores brutos transformados em estados de alto nível antes de alimentar serviços inteligentes. O mesmo princípio aplica-se ao CTI: modelos de ML podem classificar indicadores, inferir provável trajetória do ataque e atribuir escores de prioridade, desde que recebam entradas já enriquecidas e padronizadas.

Em suma, o enriquecimento contínuo—socio-organizacional, semântico e analítico—é imprescindível para manter precisão, relevância e tempestividade da Inteligência de Ameaças, permitindo respostas proativas, redução de falsos-positivos e defesas cibernéticas mais robustas em face de um panorama de ameaças em constante mutação.(41, 44)

2.2 TRABALHOS CORRELATOS

Vários enfoques para melhorar o enriquecimento de CTI (Cyber Threat Intelligence) com base em OSINT (Open Source Intelligence) têm sido propostos em diversos estudos. Nesta seção, revisamos as abordagens descritas nesses trabalhos e as comparamos com nossa metodologia.

O estudo [16] discute a **Enhanced Threat Intelligence Platform** (ETIP), que aprimora as Plataformas de Inteligência de Ameaças (TIPs) no processamento de dados provenientes de OSINT. O ETIP integra informações sobre ameaças cibernéticas com dados de infraestrutura, melhorando a análise, visualização e compartilhamento de informações.

Em [38], é apresentado um sistema que automatiza a extração e categorização de CTI usando dados de OSINT. A abordagem proposta emprega redes neurais convolucionais e análise de dependência sintática para superar as limitações dos métodos tradicionais, identificando novos *Indicators of Compromise* (IoCs) com uma precisão superior a 84% e classificando-os com eficiência de 94%. Segundo os autores, essa metodologia enriquece significativamente os dados de CTI, melhorando a geração de alertas e fortalecendo as medidas de segurança por meio de uma análise detalhada do panorama de ameaças.

Complementarmente, a relevância do enriquecimento de dados de CTI por meio do uso de OSINT é explorada em [15]. Esta pesquisa destaca a importância de integrar dados de fontes públicas e restritas para compreender plenamente as ameaças cibernéticas em evolução. Utilizando técnicas avançadas de mineração de dados e aprendizado de máquina, o enriquecimento de CTI ajuda as organizações a responder de forma mais proativa, aprimorando sua capacidade de prever e mitigar ataques cibernéticos. O estudo defende uma abordagem holística à inteligência de ameaças, enfatizando a qualidade e a integração eficiente de dados no processo de enriquecimento.

Por outro lado, [13] aborda o enriquecimento de CTI com OSINT para fortalecer políticas de segurança em infraestruturas críticas. Os autores sugerem que a integração de OSINT adiciona uma camada contextual ao CTI, levando a uma avaliação de riscos mais abrangente e a uma detecção de ameaças mais precisa. O estudo propõe uma arquitetura funcional que utiliza políticas de segurança baseadas em contexto e risco, com CTI enriquecido por OSINT permitindo ajustes dinâmicos nas medidas de segurança. Esse modelo melhora a integração entre plataformas de compartilhamento de inteligência e sistemas de controle de acesso, resolvendo limitações atuais em revogar acessos quando as condições das políticas mudam, resultando em uma gestão de segurança mais resiliente contra ameaças emergentes.

Questões relacionadas à integração de OSINT em práticas de CTI em diversas organizações são discutidas em [14], destacando a importância de estruturar adequadamente a coleta e análise de informações de fonte aberta para aprimorar a postura de cibersegurança. Por meio de uma Revisão de Literatura Semi-sistemática e entrevistas com especialistas, o estudo propõe um novo quadro conceitual que orienta as organizações a utilizar OSINT de forma eficaz para enfrentar ameaças cibernéticas de maneira proativa e informada.

A pesquisa de Arazzi et al. (45) apresenta uma análise aprofundada das técnicas de Processamento de Linguagem Natural (NLP) aplicadas à coleta e análise de dados de ameaças cibernéticas, incluindo informações de mídias sociais, Clear Web e Dark/Deep Web. O estudo destaca a importância da extração de entidades nomeadas (NER) e da identificação de eventos para estruturar dados não formatados, além de explorar a construção de grafos de conhecimento que representam as relações entre diferentes elementos de cibersegurança. Essa abordagem oferece um panorama detalhado de como as ferramentas de NLP podem automatizar a identificação, classificação e correlação de informações relevantes, como TTPs e IoCs, a partir de diversas fontes textuais. A integração dessas técnicas, em especial as baseadas em Large Language Models (LLMs), pode aprimorar significativamente a capacidade de converter grandes volumes de dados brutos de OSINT em inteligência acionável e contextualizada, fornecendo um suporte robusto para a fase de "Análise & Correlação".

Por fim, [43] apresenta um esquema extensível de enriquecimento de dados projetado para melhorar a implementação de serviços inteligentes em ambientes de Internet das Coisas (IoT). Embora o documento

não trate especificamente do enriquecimento de dados para CTI com OSINT, as metodologias propostas podem ser adaptadas para esse contexto, visando à segurança e eficiência na aquisição e análise de dados abertos e conectados. A aplicação dessas técnicas pode aprimorar a coleta de dados de inteligência, representando um avanço significativo no campo da cibersegurança e da inteligência de ameaças.

Tabela 2.1: Resumo dos Estudos de Enriquecimento de CTI Utilizando OSINT

Estudo	**Objetivo**	**Abordagem**	**Limitações**
Gonzalez Granadillo et al.(2021)	Aprimorar TIPs com integração de OSINT.	Integra dados de ameaças e infraestrutura.	Foco em TIPs, carece de métodos de classificação de CTI.
Zhao et al.(2020)	Automatizar a categorização de CTI utilizando OSINT.	Utiliza CNNs e análise sintática.	Limitado a fontes de OSINT.
Sun et al.(2023)	Enriquecer CTI com dados públicos e restritos.	Mineração de dados e aprendizado de máquina.	Falta de detalhes sobre automação.
Osliak et al.(2023)	Melhorar políticas de segurança com OSINT.	Arquitetura baseada em contexto e risco.	Problemas com integração dinâmica de políticas.
Slinde (2021)	Estruturar OSINT para cibersegurança.	Revisão de literatura e contribuições de especialistas.	Carece de validação em cenários reais.
Park et al.(2021)	Melhorar serviços de IoT via enriquecimento de dados.	Esquema extensível de dados.	Não é focado diretamente em CTI.
Arazzi et al. (2025)	Revisar e analisar técnicas de NLP aplicadas à CTI.	Explora métodos de extração automática de indicadores e entidades a partir de grandes volumes de texto usando NLP.	Não propõe pipeline ou arquitetura de enriquecimento, foca em revisão e análise de técnicas.

Com base nos estudos comparados na Tabela 2.1, a lacuna identificada refere-se à necessidade de uma metodologia mais estruturada e específica para o enriquecimento de CTI com OSINT, que ofereça: Integração detalhada e eficiente de múltiplas fontes de dados; e Aplicação prática e dinâmica na adaptação de políticas de segurança em tempo real, aspecto ainda não profundamente abordado. A Tabela 2.2 sumariza os enfoques dos trabalhos correlatos, contrastando-os com a metodologia proposta nesta pesquisa.

A análise dos trabalhos revisados (Tabela 2.1) mostra que, apesar das contribuições relevantes, nenhum deles equilibra de forma consistente automação, contextualização e integração de múltiplas fontes OSINT. Em geral, observa-se que alguns estudos priorizam a automação, mas com escopo restrito de dados; outros oferecem boa capacidade de contextualização, porém dependem fortemente de processamento manual; e há aqueles que integram diferentes fontes, mas sem um método estruturado para correlacionar e organizar as informações de forma acionável. O framework proposto neste trabalho se distingue ao combinar, em um processo metodológico claro e replicável, a contextualização aprofundada dos indicadores com a integração coerente de múltiplas fontes OSINT, utilizando o modelo 5W3H como estrutura central. Embora a automação não seja intrínseca ao framework, sua implementação na ferramenta desenvolvida nesta pesquisa permite operacionalizar as etapas propostas, aumentando a eficiência e favorecendo a adaptação a diferentes cenários e políticas de segurança.

Tabela 2.2: Comparação entre os trabalhos correlatos e o presente estudo quanto ao uso de CTI, OSINT, NLP e Enriquecimento

Trabalho	**CTI**	**OSINT**	**Enriquecimento**
Arazzi et al. (2025)	✓	✓	–
Park et al. (2021)	–	–	✓
Osliak et al. (2023)	✓	–	–
Sun et al. (2023)	✓	✓	–
González-Granadillo et al. (2021)	✓	✓	✓
Slinde (2023)	✓	✓	–
Zhao et al. (2020)	✓	✓	✓
Este trabalho	✓	✓	✓

Para viabilizar esses diferenciais na prática, a metodologia estrutura um conjunto de perguntas norteadoras que orientam a identificação, extração e categorização de dados relevantes, organizando-os em um fluxo lógico capaz de alimentar plataformas de inteligência compartilhada. Essa abordagem permite que as organizações ajustem políticas de segurança em tempo real, fortaleçam a resiliência contra ameaças emergentes e adotem uma postura mais proativa no tratamento de incidentes.

3 DISCUSSÃO DO PROBLEMA E PROPOSTA

O avanço contínuo das ameaças cibernéticas exige que a Inteligência de Ameaças Cibernéticas (CTI) vá além da mera coleta e catalogação de indicadores técnicos isolados, como hashes, IPs e domínios. O contexto operacional moderno demanda que a inteligência entregue não apenas sinais, mas significado: o verdadeiro valor da CTI está em sua capacidade de fornecer contexto qualificado, apoiar decisões estratégicas e permitir respostas rápidas e eficazes diante de cenários complexos e dinâmicos. Nesse sentido, o enriquecimento de dados, especialmente com informações oriundas de fontes abertas (OSINT), torna-se fundamental para transformar dados fragmentados em conhecimento aprofundado, agregando perspectiva temporal, geográfica, motivacional e técnica ao que, de outra forma, seriam apenas alertas desconexos.

Entretanto, é preciso ressaltar que enriquecer por enriquecer pode ser contraproducente. A inclusão indiscriminada de volumes excessivos de informações pode sobrecarregar equipes, gerar ruído e comprometer a qualidade analítica, dificultando a distinção entre o que é relevante e o que é apenas acessório. Portanto, o enriquecimento precisa ser guiado por critérios claros e alinhados aos objetivos da defesa cibernética, selecionando e integrando apenas dados que ampliem a utilidade, precisão e acionabilidade da CTI. Assim, o processo de enriquecimento deixa de ser uma simples acumulação de dados e passa a ser uma etapa estratégica para elevar a inteligência ao patamar necessário para subsidiar decisões críticas e fortalecer a resiliência organizacional.

O processo de enriquecimento da CTI tem sido frequentemente executado de forma ad hoc, fragmentada ou excessivamente dependente de ferramentas automatizadas. Essa dependência pode resultar na perda de nuances importantes e na geração de "ruído"analítico, ou seja, informações que não agregam valor ou que são de difícil interpretação. A "Pirâmide da Dor"(Figura 3.1) ilustra a hierarquia da informação de ameaças, onde indicadores de baixo nível (como hashes e IPs) são fáceis de coletar, mas oferecem pouco valor e são rapidamente substituíveis pelos adversários. À medida que se sobe na pirâmide, a dificuldade de coleta aumenta (ex: TTPs, estratégias de grupos de ameaça), mas o valor e o tempo de vida da inteligência se elevam drasticamente, tornando as defesas mais resilientes. A falta de uma metodologia estruturada para o enriquecimento dificulta a progressão para níveis mais altos e valiosos da pirâmide, mantendo as organizações presas à defesa reativa contra indicadores de baixo impacto e curta duração.

Para superar essas limitações, a validação das fontes de dados, sejam elas estruturadas (como feeds de CTI) ou não estruturadas (como publicações em mídias sociais e fóruns), é um desafio crítico. A confiabilidade e a precisão das informações podem variar significativamente, e a proliferação de dados falsos ou enviesados em fontes abertas pode comprometer a qualidade da inteligência produzida. Ferramentas automatizadas auxiliam na coleta, mas o julgamento humano e a aplicação de um framework analítico são indispensáveis para verificar, validar e filtrar os dados, garantindo sua relevância e utilidade estratégica. Este trabalho propõe um framework metodológico estruturado para guiar esse processo de enriquecimento e validação, convertendo informações dispersas e superficiais em inteligência contextualizada e robusta, capaz de sustentar uma postura de cibersegurança mais proativa e adaptativa.

Figura 3.1: Pirâmide da Dor - Pyramid of Pain (2)

3.1 METODOLOGIA

A metodologia adotada neste trabalho foi concebida com o objetivo de propor e validar um framework para o enriquecimento de CTI por meio da integração sistemática de dados oriundos de OSINT. Procuramos descrever, de forma sistemática e reprodutível, o percurso metodológico seguido para conceber o framework EnricherV2, fundamentado na adaptação do modelo 5W3H, e implementar a ferramenta de enriquecimento automático que o materializa. A abordagem adotada é de natureza aplicada, articulando princípios de Design Science Research (DSR) (46) voltados à construção de um artefato inovador e útil, juntamente com experimentação controlada para avaliação do desempenho da solução. Com isso, busca-se oferecer às organizações um método prático e validado para aumentar a capacidade de contextualização e resposta diante de ameaças cibernéticas emergentes.

A primeira etapa da metodologia consistiu em uma revisão sistemática da literatura e análise comparativa de métodos existentes, com o intuito de identificar lacunas nos modelos atuais de enriquecimento de CTI com OSINT. Em seguida, foi concebido um framework teórico utilizando o critério 5W3H, estruturando as informações em torno de oito dimensões analíticas: What, Where, Who, When, Why, How, How Much e How Long. Essa estrutura foi selecionada por sua capacidade de organizar dados complexos em um modelo compreensível, coerente e passível de aplicação prática.

Com base nesse modelo, foi desenvolvida uma ferramenta computacional dedicada, projetada para aplicar a lógica do framework proposto e operacionalizar o processo de coleta e análise de dados OSINT. A ferramenta permite integrar automaticamente múltiplas fontes de dados públicas, extraindo informações relevantes e organizando-as segundo os parâmetros do 5W3H. Isso proporcionou uma coleta sistemática,

eficiente e reprodutível, aumentando a escalabilidade da metodologia.

A ferramenta foi empregada em um estudo de caso focado no enriquecimento de um Indicador de Comprometimento (IoC), especificamente um endereço IP. Por meio dela, foi possível agregar dados de diversas fontes, como AbuseIPDB, AlienVault OTX, VirusTotal, Censys, entre outras, e estruturá-los de maneira contextualizada. Esse processo resultou em uma visão mais abrangente e informativa da ameaça, demonstrando a efetividade da metodologia na prática.

Essa abordagem oferece uma alternativa flexível e acessível às soluções comerciais de alto custo baseadas em inteligência artificial e automação avançada, sendo especialmente útil para organizações que buscam soluções personalizadas de CTI. No capítulo seguinte, serão apresentados os detalhes técnicos da implementação da ferramenta, bem como os resultados obtidos na sua aplicação prática.

3.2 PROPOSTA

O EnricherV2 é um framework proposto para padronizar e potencializar o processo de enriquecimento de dados na Inteligência de Ameaças Cibernéticas (CTI), com ênfase na integração de informações provenientes de fontes abertas (OSINT). Diferentemente de abordagens tradicionais centradas apenas em coleta e catalogação de indicadores, o EnricherV2 propõe uma metodologia iterativa, multidimensional e adaptável, capaz de transformar dados brutos, frequentemente dispersos, em inteligência acionável, robusta e contextualizada para o cenário de defesa cibernética. Este framework faz uso adaptado do modelo 5W3H, organizando o processo de enriquecimento em oito dimensões fundamentais para transformar informações dispersas ou superficiais em inteligência contextualizada, robusta e operacional.

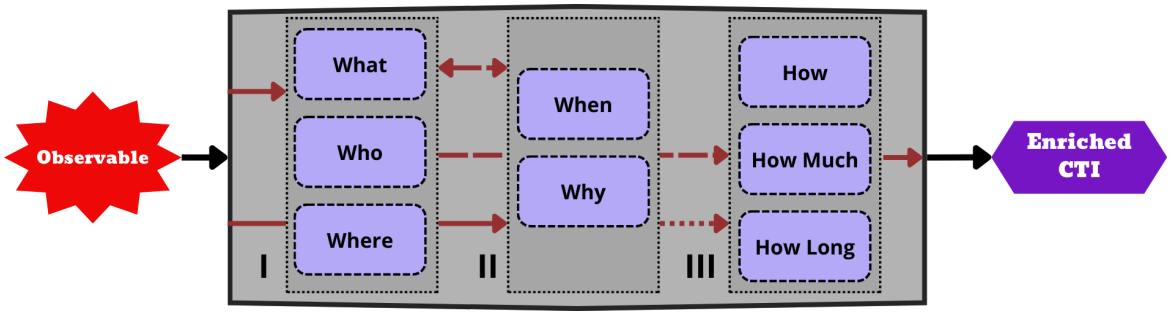

Figura 3.2: Framework EnricherV2

3.2.1 5W3H

O 5W3H empregado no framework EnricherV2 foi proposto por (36) para avaliação de plataformas e padrões de CTI . No contexto daquele trabalho, o 5W3H auxilia a mapear de forma holística aspectos fundamentais de uma ameaça cibernética — o que, onde, quando, quem, por que, como, quanto e por quanto tempo. Isso facilita a análise da completude, aplicabilidade e interoperabilidade das plataformas e padrões avaliados.

Ao trazer esse conceito para o enriquecimento de CTI com OSINT, o EnricherV2 amplia a utilidade do

5W3H. Não se limita a avaliar ferramentas, mas propõe um roteiro analítico para guiar a busca, correlação e validação de dados em todas as etapas do enriquecimento de inteligência de ameaças. O uso do 5W3H permite ao analista estruturar perguntas que organizam e potencializam a contextualização dos dados coletados, facilitando a transformação de informações dispersas em inteligência robusta e acionável. Assim, a metodologia não só mensura a capacidade das plataformas, mas fornece um framework operacional para aumentar a qualidade, relevância e aplicabilidade dos dados de CTI obtidos de fontes abertas.

Em resumo: o 5W3H, já validado para avaliação de plataformas, se revela igualmente valioso como núcleo de uma abordagem sistemática para enriquecimento de inteligência — alinhando o framework EnricherV2 com as melhores práticas internacionais de produção de CTI.

A aplicação do framework ao processo de enriquecimento de CTI é realizada de forma estruturada e iterativa, proporcionando um avanço progressivo na compreensão da ameaça a partir de múltiplas dimensões contextuais. O fluxo sugerido para utilização do modelo parte da priorização das perguntas “What”, “Who” e “Where”, visando rapidamente estabelecer um conjunto mínimo de dados essenciais sobre a ameaça: qual é o evento (o quê), quem está envolvido (quem) e quais as origens ou alvos (onde). Esses parâmetros básicos servem como ponto de partida para guiar buscas complementares e criar uma visão inicial do incidente.

Na sequência, a análise aprofunda-se nas perguntas “When” e “Why”, que auxiliam a reconstruir cronologias relevantes e compreender motivações ou objetivos dos agentes maliciosos. O conhecimento sobre quando as atividades ocorreram permite identificar relações temporais, campanhas simultâneas ou tendências, enquanto entender o “porquê” revela possíveis interesses estratégicos, ideológicos ou financeiros.

Por fim, são abordadas as dimensões “How”, “How Much” e “How Long”, que tratam dos métodos e técnicas empregadas pelo adversário, da extensão dos danos causados e da duração ou persistência da ameaça. Essas informações são determinantes para orientar ações de resposta técnica, estimar recursos necessários para mitigação e aprimorar os controles de defesa.

É importante destacar que, apesar da ordem sugerida, o framework foi concebido para ser flexível e adaptativo: em muitos casos, as informações disponíveis podem não seguir a sequência ideal ou determinadas respostas podem já estar parcialmente conhecidas no início da análise. Dessa forma, o ciclo 5W3H deve ser entendido como um guia estruturante e não como um roteiro rígido. O analista pode iniciar por qualquer dimensão conforme a natureza dos dados acessíveis, retomando etapas anteriores sempre que novas evidências surgirem ou lacunas forem identificadas.

Essa abordagem iterativa e adaptável permite a reavaliação contínua das informações, favorecendo a construção de uma inteligência contextualizada, robusta e confiável. O processo de enriquecimento, portanto, evolui até que todas as questões relevantes sejam respondidas com o grau de detalhamento necessário para subsidiar decisões e fortalecer as medidas de defesa cibernética.

Tabela 3.1: Parâmetros 5W3H

Parâmetro	Itens
What	Tipos de ameaça, Natureza técnica, Eventos específicos
Who	Atores envolvidos, Principais agentes, Vítimas
Where	Origem geográfica, Sistemas-alvo, Trajetória do ataque
When	Momento das atividades, Linha do tempo de detecção, Linha do tempo do ataque
Why	Motivações, Seleção de alvos, Justificativa do momento
How	Técnicas utilizadas, Acesso inicial, Persistência
How Much	Impacto, Infraestrutura afetada, Recursos para recuperação
How Long	Duração, Tempo até a detecção, Tempo até a recuperação

3.2.1.1 What

No contexto da inteligência de ameaças cibernéticas, o parâmetro What refere-se à identificação clara e objetiva do evento ou incidente analisado. Trata-se de determinar com precisão qual tipo de ameaça está em curso, incluindo a classificação do ataque — como malware, phishing, ransomware, brute force, exploração de vulnerabilidades, entre outros — e a caracterização dos principais artefatos técnicos envolvidos, como endereços IP, domínios, arquivos ou URLs maliciosos. A delimitação do "What" constitui a base factual do processo de enriquecimento e influencia diretamente a qualidade das etapas subsequentes de análise.

A correta definição desse parâmetro é fundamental para o direcionamento eficiente das ações de resposta e a priorização de recursos de defesa. Quando o "What" é identificado com precisão, torna-se possível buscar, de maneira estruturada, informações correlatas em diferentes bases OSINT, correlacionar eventos dispersos e identificar padrões emergentes que poderiam passar despercebidos em uma abordagem menos sistemática. O enriquecimento desse aspecto, portanto, não apenas aprimora a contextualização do incidente, mas também contribui para a detecção precoce de campanhas coordenadas e o fortalecimento do processo decisório nas equipes de segurança.

Um exemplo prático desse processo pode ser observado na identificação de um endereço IP reportado em múltiplos feeds de abuso devido a tentativas reiteradas de ataques de força bruta. Esse cenário permite classificar a ocorrência como parte de uma atividade automatizada de botnet, o que, por sua vez, orienta bloqueios e ajustes de políticas preventivas. De modo semelhante, ao analisar um arquivo executável suspeito disseminado via e-mail, pode-se identificar rapidamente se se trata de uma amostra de ransomware ou de um trojan bancário, possibilitando resposta imediata e comunicação assertiva com os públicos internos afetados.

Como ilustração, imagine um analista que detecta um domínio recém-registrado e associado a uma onda de e-mails fraudulentos direcionados a clientes de uma instituição financeira. Consultando plataformas OSINT, como VirusTotal e AbuseIPDB, o analista verifica que aquele domínio já figura em diversos

relatos de phishing, reforçando sua classificação como ameaça ativa. O esclarecimento do “What”, neste caso, sustenta a tomada de decisão para bloqueio do domínio e a emissão de alertas, evitando possíveis impactos à organização e aos seus usuários.

3.2.1.2 Who

O parâmetro "Who"refere-se à identificação dos atores responsáveis pelas ações maliciosas, bem como à compreensão de suas afiliações, hierarquias e possíveis alianças. Essa dimensão abrange desde grupos de cibercriminosos organizados, insiders e hacktivistas, até atores patrocinados por Estados-nação, exigindo uma abordagem multifacetada para caracterizar corretamente os responsáveis por uma ameaça. A definição do “Who” envolve a coleta e correlação de dados provenientes de fontes abertas, como o monitoramento de mídias sociais, acompanhamento de fóruns na dark web e análise de comunicações vazadas ou exposições de identidade em ambientes clandestinos.

A correta identificação dos agentes envolvidos possibilita uma análise mais apurada das intenções, capacidades técnicas e estratégias adotadas pelos adversários. O enriquecimento desse parâmetro permite diferenciar ataques oportunistas de campanhas direcionadas, identificar padrões de comportamento recorrentes e antecipar movimentos de grupos conhecidos, além de subsidiar ações coordenadas com outras organizações ou órgãos governamentais. Compreender a estrutura interna, as hierarquias e as parcerias estabelecidas dentro de um grupo de ameaça também contribui para a previsão de possíveis alvos e para o desenvolvimento de contramedidas mais eficazes e alinhadas ao perfil do atacante.

Como exemplo prático, considere o caso em que um grupo de ransomware torna pública, em um fórum da dark web, uma lista de organizações comprometidas. Ao monitorar essas postagens e analisar a linguagem empregada, as táticas relatadas e os padrões de negociação, um analista de CTI pode atribuir a campanha a um grupo específico já conhecido no ecossistema de ameaças. Esse entendimento do “Who” permite antecipar potenciais vítimas, compartilhar alertas com o setor afetado e ajustar estratégias defensivas para mitigar riscos associados àquele agente.

3.2.1.3 Where

No contexto da inteligência de ameaças cibernéticas, o parâmetro Where diz respeito à identificação da origem geográfica e dos alvos específicos envolvidos em um incidente de segurança. Essa dimensão abrange tanto a localização do ponto inicial do ataque — como o país, a região ou o provedor de onde partiu a ação maliciosa — quanto o mapeamento dos destinos afetados dentro da rede organizacional. A definição do “Where” envolve a análise de dados de geolocalização de endereços IP, informações obtidas por meio de consultas WHOIS, e o monitoramento de servidores de Comando e Controle (C2) que possam estar sendo utilizados para coordenação ou exfiltração de dados.

A correta determinação desse parâmetro é fundamental para subsidiar ações de contenção e para o entendimento das rotas percorridas pelo atacante, desde a entrada inicial até eventuais movimentos laterais realizados na infraestrutura interna. O enriquecimento do “Where” possibilita ainda o reconhecimento de padrões geográficos recorrentes, a identificação de vulnerabilidades exploradas em segmentos específicos

da rede e a avaliação de riscos na cadeia de suprimentos, especialmente em ambientes que dependem de parceiros externos ou fornecedores. Esse mapeamento detalhado fortalece a consciência situacional da equipe de resposta, permitindo a priorização de ativos críticos e a adoção de medidas preventivas ajustadas ao contexto de exposição.

Como exemplo prático, pode-se citar o rastreamento de uma campanha de spear phishing originada em um país estrangeiro, cujo vetor inicial foi identificado por meio de logs de firewall associados a um bloco de IPs conhecido por hospedar atividades maliciosas. Ao analisar a trajetória do ataque dentro da rede, verificou-se que o atacante obteve acesso a um servidor de arquivos sensível e tentou realizar movimentos laterais para outros departamentos, evidenciando a importância de monitorar não apenas o ponto de entrada, mas todo o percurso da ameaça. Esse entendimento do “Where” permitiu à organização reforçar controles nos pontos comprometidos e aprimorar procedimentos de monitoramento e resposta em áreas estratégicas.

3.2.1.4 When

O parâmetro "When"diz respeito à determinação precisa do momento em que as atividades maliciosas ocorrem, abrangendo desde o planejamento e preparação até a execução efetiva dos ataques. Essa dimensão envolve o estabelecimento de uma linha do tempo detalhada dos eventos, permitindo que analistas compreendam não apenas quando um indicador de compromisso foi detectado, mas também a relação temporal entre diferentes etapas de uma campanha. O uso de OSINT possibilita a identificação de discussões preliminares em fóruns de hackers, o monitoramento do surgimento e disseminação de IoCs e a análise da evolução de campanhas em ambientes digitais.

A correta avaliação do “When” é fundamental para a reconstrução cronológica de incidentes e para a detecção de padrões de atividade maliciosa, como picos em determinados horários, datas comemorativas ou períodos de instabilidade geopolítica e financeira. Esse entendimento temporal contribui para a antecipação de ameaças, permitindo que organizações preparem suas defesas em momentos críticos, além de apoiar investigações retrospectivas e análises de tendência para futuras ocorrências.

Como exemplo, pode-se citar o monitoramento de campanhas de phishing que intensificam suas atividades durante períodos eleitorais ou grandes eventos esportivos. Ao identificar, por meio de logs, feeds públicos e discussões em comunidades online, que determinados ataques ocorrem em sincronia com fatos relevantes do cenário externo, é possível antecipar movimentações adversas e adotar medidas proativas de proteção. Dessa forma, a dimensão “When” viabiliza uma resposta mais ágil e estratégica diante do cenário dinâmico das ameaças cibernéticas.

3.2.1.5 Why

Motivações muitas vezes definem o perfil e o alcance de uma ameaça cibernética. O parâmetro Why busca compreender as razões subjacentes que levam atacantes a selecionar determinados alvos e executar suas ações, sejam elas de cunho ideológico, político, financeiro ou mesmo de demonstração de capacidade técnica. A análise dessas intenções vai além da superfície técnica do incidente e exige uma investigação cuidadosa dos contextos sociais e econômicos em que o ataque ocorre.

Ferramentas OSINT desempenham papel crucial nesse processo ao permitir o monitoramento de registros públicos, discussões em mídias sociais, fóruns clandestinos e comunicados de grupos na dark web. Por meio da coleta e análise dessas informações, é possível identificar justificativas explícitas apresentadas pelos próprios atores, sinais de retaliação, tentativas de extorsão ou participação em campanhas hacktivistas. Essa camada de entendimento auxilia as organizações não apenas a adaptar suas defesas de forma mais precisa, mas também a antecipar possíveis movimentos e alinhar estratégias de resposta aos objetivos percebidos dos adversários.

Como ilustração, considere uma série de ataques direcionados a instituições de saúde acompanhada por publicações de um grupo hacktivista reivindicando motivação política e crítica ao sistema de saúde local. A investigação do "Why", nesse caso, permite que a organização compreenda o contexto maior por trás do incidente, preparando-se para novas ondas de ataques semelhantes e ajustando sua comunicação estratégica com stakeholders internos e externos.

3.2.1.6 How

Compreender de que maneira uma ameaça se concretiza é essencial para a eficácia da defesa cibernética. O parâmetro How busca desvendar os métodos, técnicas e ferramentas empregadas pelos agentes maliciosos em suas campanhas, desde o vetor inicial de acesso até as etapas subsequentes de persistência e movimentação lateral nas redes. Essa investigação vai além da simples identificação do ataque, detalhando como os adversários exploram vulnerabilidades, utilizam kits de malware, executam campanhas de phishing ou implementam mecanismos de evasão e exfiltração de dados.

A análise desse parâmetro é enriquecida por informações extraídas de OSINT, como o monitoramento de fóruns especializados, discussões na dark web e relatos de incidentes públicos. Ao rastrear a evolução de malwares, examinar tutoriais compartilhados por atores maliciosos e correlacionar padrões de ataque, torna-se possível mapear as TTPs (Táticas, Técnicas e Procedimentos) mais frequentes. Esse conhecimento orienta a adaptação de controles defensivos, a implementação de medidas preventivas específicas e o desenvolvimento de planos de resposta alinhados ao cenário de ameaças vigente.

Por exemplo, ao analisar uma onda de ataques de ransomware que utiliza phishing como vetor inicial, a investigação do "How" permite identificar as etapas seguidas pelos atacantes — desde o envio de e-mails fraudulentos até a execução de scripts automatizados para criptografia de arquivos e solicitação de resgate. Esse entendimento detalhado possibilita não apenas uma resposta mais eficaz ao incidente, mas também a antecipação de variantes e o fortalecimento contínuo das defesas organizacionais.

3.2.1.7 How Much

Avaliar o real impacto de uma ameaça é um dos grandes desafios da inteligência de ameaças cibernéticas. O parâmetro How Much concentra-se na quantificação dos danos causados por um incidente, considerando não apenas as perdas materiais e financeiras, mas também as consequências reputacionais, operacionais e legais para a organização afetada. Essa análise abrange a extensão da interrupção dos serviços, o número de sistemas comprometidos, os custos estimados de remediação e o tempo necessário para

recuperação total.

O enriquecimento desse parâmetro com dados de OSINT envolve o monitoramento de relatórios públicos de incidentes, notícias de mídia especializada e discussões em redes sociais, que frequentemente trazem informações sobre prejuízos experimentados por outras organizações diante de ameaças semelhantes. Ao coletar e comparar esses relatos, torna-se possível medir as implicações mais amplas de um ataque, incluindo o alcance de vazamentos de dados, multas aplicadas por órgãos reguladores ou impactos indiretos sobre a confiança de clientes e parceiros.

Por exemplo, em uma situação de ataque de ransomware que paralisa uma rede hospitalar, o "How Much" é ilustrado pelo cálculo dos dias de indisponibilidade dos sistemas, pela estimativa dos valores de resgate exigidos, pelos custos de restauração de backups e pelo volume de dados potencialmente comprometidos. Essa avaliação abrangente permite que a organização planeje não só a resposta imediata ao incidente, mas também estratégias para mitigar efeitos de longo prazo e fortalecer sua resiliência diante de futuras ameaças.

3.2.1.8 How Long

Compreender por quanto tempo uma ameaça permanece ativa é fundamental para a avaliação do risco e o aprimoramento das estratégias de defesa. O parâmetro How Long se dedica à análise da duração e persistência de uma ameaça, examinando o período entre o início da atividade maliciosa, sua detecção, a resposta efetiva e a eventual erradicação. Essa dimensão abrange tanto ameaças de curta duração, como ataques pontuais de phishing, quanto campanhas prolongadas de comprometimento, em que o adversário mantém acesso à infraestrutura da vítima por semanas ou até meses.

A coleta de dados por meio de OSINT contribui significativamente para esse entendimento, ao permitir o acompanhamento de registros históricos de incidentes, alertas emitidos em tempo real e discussões contínuas em comunidades especializadas. Ao cruzar informações sobre o tempo de exposição, padrões de reinfecção e relatos de recorrência em setores específicos, é possível estimar não apenas a janela de atuação do atacante, mas também a rapidez e a eficiência dos processos de recuperação implementados pela organização.

Por exemplo, ao investigar uma campanha de infecção por malware que permaneceu indetectada durante vários meses em múltiplas empresas do setor financeiro, o parâmetro "How Long" evidencia a necessidade de mecanismos de detecção mais avançados e de monitoramento contínuo após a remediação inicial. Essa análise temporal orienta ajustes em planos de resposta, reforça práticas de lições aprendidas e apoia a avaliação do risco de recorrência, tornando a postura de defesa mais proativa e resiliente diante de ameaças persistentes.

4 RECURSOS E MÉTODOS

Este capítulo apresenta os principais recursos e métodos utilizados para a implementação do framework EnricherV2, abrangendo as ferramentas, plataformas, feeds de dados e ambientes empregados na coleta e enriquecimento da inteligência de ameaças cibernéticas. Além de descrever detalhadamente as tecnologias e fontes adotadas, o capítulo realiza uma comparação do EnricherV2 com abordagens tradicionais e recentes de enriquecimento de CTI, destacando as diferenças, vantagens e limitações do processo proposto em relação aos métodos já estabelecidos na literatura.

4.1 FONTES DE DADOS UTILIZADAS NA PROPOSTA

Para a aplicação da metodologia proposta no contexto de enriquecimento de CTI utilizando OSINT, diversas fontes de dados públicas e abertas foram utilizadas. A seleção dessas fontes foi baseada em sua relevância, credibilidade e capacidade de fornecer informações contextuais e acionáveis sobre ameaças cibernéticas.

Abaixo estão as principais fontes de dados utilizadas durante o estudo de caso e na implementação do EnricherV2, juntamente com uma justificativa para sua seleção e uma explicação de como cada uma contribuiu para a análise. É importante observar que os conjuntos de dados utilizados neste estudo foram selecionados com base em sua relevância para os objetivos propostos. No entanto, a metodologia foi projetada para evitar limitações ou restrições a fontes específicas, mitigando o risco de dependência exclusiva de um único conjunto de dados. O foco principal é orientar o processo de enriquecimento das informações de forma robusta e flexível.

A seleção das fontes de dados foi guiada por diversos fatores. Todas as fontes são amplamente reconhecidas pela comunidade de segurança cibernética e têm um longo histórico de fornecimento de dados precisos e úteis para coleta de dados e análise de ameaças. A combinação de diferentes tipos de fontes (análise de malware, dispositivos conectados, phishing, inteligência compartilhada e redes sociais) forneceu uma visão abrangente e contextual do IoC alvo. Para a aplicação e avaliação da metodologia proposta neste estudo, foi desenvolvida, como dito anteriormente, o EnricherV2. Esta ferramenta tem como objetivo principal facilitar a coleta, organização e análise preliminar de informações das diversas fontes de dados OSINT listadas a seguir, alinhando-se com os parâmetros do framework 5W3H. Enquanto as fontes de dados são selecionadas por sua relevância e credibilidade, a ferramenta de software desenvolvida agiliza o processo de interação com estas fontes. A aplicação prática desta ferramenta no enriquecimento de CTI é demonstrada no Capítulo 5: RESULTADOS EXPERIMENTAIS. É importante ressaltar que, embora a ferramenta automatize etapas da coleta, estruturação e análise inicial, a análise crítica e o enriquecimento contextual final permanecem como responsabilidade do analista.

4.1.1 FERRAMENTAS DE COLETA

As ferramentas de coleta apresentadas nesta subseção compõem um conjunto de fontes abertas, bases públicas e plataformas colaborativas amplamente utilizadas para levantamento de dados relacionados a ameaças cibernéticas. Esses recursos permitem identificar reputação, histórico de abuso, geolocalização, informações de registro e características técnicas de endereços IP e domínios investigados. A partir dessas plataformas, são extraídos indicadores iniciais — como denúncias de abuso, participação em botnets, envolvimento com spam ou anonimato via TOR — que fundamentam as etapas posteriores de análise e contextualização da ameaça.

4.1.1.1 AbuseIPDB

AbuseIPDB é uma plataforma que fornece um repositório de endereços IP reportados associados a atividades maliciosas, como tentativas de hackeamento, spams e outros crimes cibernéticos.

URL: <https://www.abuseipdb.com>

Uso: Usado para verificar a reputação do IP alvo com base nos casos de abuso reportados.

4.1.1.2 AlienVault Open Threat Exchange (OTX)

OTX é uma plataforma colaborativa de inteligência de ameaças que fornece feeds em tempo real de indicadores de compromisso (IoCs), dados sobre ameaças e padrões de ataque.

URL: <https://otx.alienvault.com>

Uso: Usado para verificar o IP em relação a campanhas de ameaças conhecidas e IoCs compartilhados pela comunidade de segurança.

4.1.1.3 Censys

Censys realiza varreduras e indexa dispositivos expostos na internet, fornecendo informações detalhadas sobre suas vulnerabilidades, configurações e portas abertas.

URL: <https://censys.io>

Uso: Fornece informações sobre serviços expostos e vulnerabilidades associadas ao IP alvo.

4.1.1.4 TOR Project

O TOR Project facilita a comunicação anônima na internet, roteando o tráfego por meio de uma rede descentralizada de nós.

URL: <https://www.torproject.org>

Uso: Usado para determinar se o IP alvo está vinculado a nós de saída do TOR, indicando possível uso

de anonimato.

4.1.1.5 IPInfo

IPInfo fornece informações detalhadas sobre endereços IP, incluindo geolocalização, nome de domínio e riscos associados.

URL: <https://ipinfo.io>

Uso: Usado para coletar metadados sobre o IP, como localização e ISP, para ajudar a contextualizar a ameaça.

4.1.1.6 ICANN WHOIS

A base de dados WHOIS da ICANN fornece informações sobre registrantes de domínios e a propriedade de endereços IP.

URL: <https://whois.icann.org>

Uso: Usado para coletar detalhes de registro do IP e verificar padrões ou associações suspeitas.

4.1.1.7 CleanTalk

CleanTalk é um serviço de prevenção de spam que monitora endereços IP para atividade suspeita ou maliciosa relacionada a spam.

URL: <https://cleantalk.org>

Uso: Usado para verificar se o IP está associado a spam ou outro comportamento abusivo.

4.1.1.8 GreyNoise

GreyNoise fornece dados sobre atividades de varredura em larga escala na internet e diferencia ruído de ameaças reais.

URL: <https://www.greynoise.io>

Uso: Usado para verificar se o IP alvo está envolvido em varreduras de internet em grande escala ou faz parte de um botnet.

4.1.2 FERRAMENTAS DE ANÁLISE

Na etapa de análise, as ferramentas selecionadas têm como objetivo transformar os dados coletados em inteligência contextualizada, correlacionando múltiplos pontos de informação e atribuindo significado operacional aos achados. Aqui, são utilizados frameworks, bases de conhecimento e portais especializados para mapeamento de táticas, técnicas e procedimentos (TTPs), avaliação de reputação, detecção de

comportamentos maliciosos em larga escala e validação dos indicadores identificados. Esses recursos são essenciais para interpretar as evidências, identificar padrões de ataque e subsidiar recomendações técnicas ou estratégicas para resposta a incidentes cibernéticos.

4.1.2.1 VirusTotal

VirusTotal agrega e analisa arquivos e URLs suspeitos, verificando-os em vários motores antivírus para identificar possíveis malwares ou ameaças.

URL: <https://www.virustotal.com>

Uso: Usado para verificar se o IP foi sinalizado por malware ou atividade maliciosa.

4.1.2.2 SANS Internet Storm Center

O SANS Internet Storm Center acompanha incidentes globais de segurança cibernética, fornecendo alertas em tempo real e análise sobre ameaças emergentes.

URL: <https://isc.sans.edu>

Uso: Verificado para obter relatórios e alertas relacionados ao IP alvo da comunidade SANS.

4.1.2.3 Kaspersky Threat Intelligence Portal

A plataforma de inteligência de ameaças da Kaspersky oferece insights sobre endereços IP, arquivos e URLs associados a ameaças cibernéticas.

URL: <https://opentip.kaspersky.com>

Uso: Usado para coletar inteligência sobre a reputação do IP alvo e sua associação com ameaças cibernéticas conhecidas.

4.1.2.4 MITRE ATT&CK Framework

O framework MITRE ATT&CK é uma base de conhecimento de táticas, técnicas e procedimentos (TTPs) usados por adversários em segurança cibernética.

URL: <https://attack.mitre.org>

Uso: Usado para mapear o comportamento do IP alvo para TTPs conhecidos, ajudando a identificar o tipo de ataque ou ameaça com a qual ele pode estar associado.

4.1.2.5 Cyber Kill Chain

A Cyber Kill Chain é um modelo que descreve as fases de um ataque cibernético, desde o reconhecimento inicial até a exfiltração de dados.

URL: <https://www.lockheedmartin.com/en-us/capabilities/cyber-kill-chain.html>

Uso: Usado para entender as possíveis fases do ataque associado ao IP alvo, mapeando-o para o modelo Cyber Kill Chain.

4.1.2.6 TRIAGE

TRIAGE é uma ferramenta de inteligência de ameaças projetada para coletar e analisar diversos pontos de dados sobre ameaças para obter insights acionáveis.

URL: <https://triage.io>

Uso: Usado para analisar múltiplos pontos de dados de inteligência de ameaças para obter insights mais profundos sobre as atividades do IP.

4.2 COMPARAÇÃO COM MÉTODOS EXISTENTES DE ENRIQUECIMENTO DE INTELIGÊNCIA DE AMEAÇAS

A metodologia proposta neste estudo oferece uma abordagem estruturada para o enriquecimento de CTI usando OSINT. No entanto, para contextualizar e validar sua aplicação, é importante comparar esta metodologia com outras abordagens amplamente utilizadas no campo da cibersegurança, incluindo plataformas automatizadas de inteligência de ameaças, ferramentas OSINT e métodos baseados em aprendizado de máquina e inteligência artificial (IA).

4.2.1 Plataformas Automatizadas de Inteligência de Ameaças

Plataformas automatizadas de inteligência de ameaças, como ThreatConnect, Anomali e IBM X-Force Exchange, são amplamente reconhecidas pela sua capacidade de integrar feeds OSINT com dados proprietários, correlacionar eventos e automatizar grande parte do processo de coleta e análise de dados. Essas plataformas oferecem funcionalidades avançadas, como análises preditivas e alertas automatizados, permitindo uma resposta rápida a incidentes cibernéticos.

Essas plataformas têm a vantagem de uma automação robusta, permitindo análise em tempo real, e são altamente escaláveis, capazes de lidar com grandes volumes de dados de forma eficiente. No entanto, elas têm um custo elevado, o que pode ser uma barreira para organizações menores. Além disso, tendem a ser mais fechadas e menos personalizáveis, tornando mais difícil adaptá-las a necessidades de segurança específicas.

Em comparação com plataformas totalmente automatizadas como ThreatConnect, que visam automatizar todo o ciclo de vida da inteligência de ameaças, desde a coleta de dados até a geração de alertas, a metodologia proposta, que pode ser parcialmente instrumentada pela ferramenta desenvolvida neste trabalho, oferece um enfoque diferenciado. A ferramenta desenvolvida auxilia na coleta e organização inicial de dados de diversas fontes OSINT. No entanto, o cerne da metodologia reside na aplicação estruturada

do framework 5W3H, que guia o analista na condução de uma análise contextual aprofundada e na interpretação dos dados. Embora esta abordagem, mesmo com o suporte da ferramenta, possa não atingir a mesma escalabilidade e velocidade de plataformas comerciais de grande porte, ela enfatiza a flexibilidade inerente ao processo investigativo guiado pelo analista e a capacidade de direcionar a análise para aspectos específicos da ameaça ou do IoC, conforme orientado pelo framework. Assim, a metodologia proposta, complementada pela ferramenta de apoio, representa uma alternativa acessível, especialmente para organizações com orçamentos limitados ou aquelas que priorizam uma compreensão contextual detalhada em detrimento de uma automação completa, permitindo que o analista mantenha um controle significativo sobre o fluxo e a profundidade da investigação.

Tabela 4.1: Comparação entre plataformas automatizadas e metodologia proposta

Característica	**ThreatConnect**	**Anomali**	**IBM X-Force Exchange**	**EnricherV2**
Integração de dados OSINT/proprietário	✓	✓	✓	✓
Automação completa do ciclo CTI	✓	✓	✓	–
Análise contextual estruturada	–	–	–	✓
Escalabilidade em grandes volumes	✓	✓	✓	–
Personalização/flexibilidade	–	–	–	✓
Baixo custo de implementação	–	–	–	✓
Controle do analista	–	–	–	✓

4.2.2 Ferramentas Baseadas em OSINT

Ferramentas como VirusTotal, Shodan, AbuseIPDB, Censys e GreyNoise são comumente usadas na coleta de dados para o enriquecimento de CTI, fornecendo informações valiosas sobre indicadores de comprometimento (IoCs), como endereços IP, domínios e arquivos suspeitos. Essas ferramentas são econômicas e oferecem uma variedade de fontes para análise de dados, sendo frequentemente usadas em conjunto com outras metodologias.

A Tabela 4.2 apresenta uma comparação entre diferentes ferramentas OSINT — como VirusTotal, Shodan, AbuseIPDB, Censys e GreyNoise — e a metodologia proposta, implementada no EnricherV2, evidenciando as diferenças em termos de escopo, abordagem e valor entregue ao processo de CTI. Enquanto as ferramentas listadas disponibilizam dados brutos ou parcialmente enriquecidos sobre Indicadores de Comprometimento (IoCs), o EnricherV2 aplica um modelo 5W3H expandido para integrar múltiplas fontes e gerar análise contextual estruturada, otimizando o uso operacional das informações.

A metodologia proposta complementa as ferramentas OSINT ao adicionar uma camada de análise contextual estruturada, utilizando a estrutura "What", "Where", "Who", "When", "Why", "How", "How Much"e "How Long". Enquanto ferramentas como VirusTotal e Shodan fornecem dados brutos, a metodologia proposta organiza e enriquece essas informações, tornando-as mais acionáveis e estrategicamente

relevantes. Isso permite que os analistas não apenas coletem dados, mas também obtenham uma compreensão mais profunda do contexto e das implicações de uma ameaça. Portanto, a metodologia proposta pode ser vista como uma forma de aprimorar os dados fornecidos pelas ferramentas OSINT, tornando-os mais aplicáveis para a tomada de decisões de segurança.

Tabela 4.2: Comparação entre ferramentas OSINT e metodologia proposta

Característica	**VirusTotal**	**Shodan**	**AbuseIPDB**	**Censys**	**GreyNoise**	**EnricherV2**
Dados brutos sobre IoCs	✓	✓	✓	✓	✓	✓
Gratuita/baixo custo	✓	✓	✓	✓	✓	✓
Análise contextual estruturada	–	–	–	–	–	✓
Facilidade de uso na coleta	✓	✓	✓	✓	✓	✓
Risco de sobrecarga de informação	✓	✓	✓	✓	✓	–
Valor estratégico para decisão	–	–	–	–	–	✓

4.2.3 Desenvolvimento de Ferramenta e Integração com IA

Nos últimos anos, ferramentas baseadas em aprendizado de máquina e inteligência artificial (IA) tornaram-se cada vez mais populares para o enriquecimento de CTI, particularmente em plataformas automatizadas. Ferramentas como Darktrace e CrowdStrike usam IA para detectar comportamentos anômalos e prever ameaças com base em dados históricos, permitindo respostas rápidas e preditivas a incidentes de segurança.

Soluções baseadas em IA oferecem vantagens significativas, como automação avançada e escalabilidade, permitindo que grandes volumes de dados sejam analisados rapidamente e padrões comportamentais sejam identificados com precisão. No entanto, essas soluções também enfrentam desafios, como alta complexidade e custo de implementação, além de dependerem de dados históricos para aprender e fazer previsões, o que pode ser problemático no caso de novas ameaças desconhecidas.

No que tange à integração com Inteligência Artificial (IA), a metodologia proposta, em sua concepção original do framework 5W3H, não pressupunha explicitamente o uso de IA. Contudo, no decorrer do desenvolvimento da ferramenta (EnricherV2) para apoiar a aplicação desta metodologia, foram incorporadas funcionalidades de IA em momentos específicos, particularmente para auxiliar na análise de contexto a partir dos dados coletados. Embora o nível de automação e a aplicação de IA pela ferramenta possam não ser tão extensivos quanto os de plataformas totalmente dedicadas à IA – o que ainda pode implicar certas limitações em cenários que envolvem grandes conjuntos de dados ou análise em tempo real – esta integração representa um avanço significativo em relação a uma abordagem puramente manual. A utilização de IA pela ferramenta busca oferecer uma análise contextual preliminar mais ágil e a identificação de insights relevantes, complementando a estrutura analítica do framework 5W3H sem comprometer a profundidade

e a flexibilidade investigativa que a estrutura original oferece ao analista.

A metodologia proposta, agora suportada por uma ferramenta que emprega IA pontualmente, continua a se destacar por oferecer flexibilidade e profundidade contextual no enriquecimento de CTI. Enquanto outras soluções são altamente automatizadas e podem fazer uso intensivo de IA, a abordagem deste trabalho busca um equilíbrio, configurando-se como uma opção acessível e adaptável, ideal para organizações com orçamentos limitados ou aquelas que buscam uma análise mais profunda e personalizada de suas ameaças. A incorporação de IA na ferramenta desenvolvida não tem o objetivo de replicar todas as capacidades de plataformas de IA especializadas, mas sim de aprimorar a eficiência e a capacidade analítica da metodologia 5W3H no contexto prático, enriquecendo o processo de CTI e tornando-o uma opção mais robusta e versátil.

5 RESULTADOS EXPERIMENTAIS

A presente seção apresenta e discute os resultados experimentais obtidos a partir da implementação e avaliação da metodologia proposta neste trabalho. São detalhados os procedimentos realizados, os dados coletados, bem como as principais métricas utilizadas para análise de desempenho e efetividade das soluções. Os resultados são apresentados de forma sistemática, permitindo uma avaliação crítica quanto aos objetivos traçados e às hipóteses levantadas ao longo da pesquisa.

5.1 OBJETIVO DOS EXPERIMENTOS

A etapa experimental deste trabalho teve como principal objetivo avaliar a efetividade da metodologia proposta para o enriquecimento de CTI por meio da integração sistemática de dados oriundos de OSINT. Buscou-se verificar, na prática, a capacidade do framework e da ferramenta EnricherV2 em organizar, contextualizar e ampliar significativamente as informações associadas a IoCs, proporcionando maior suporte à análise e resposta a incidentes de segurança.

Adicionalmente, buscou-se validar o desempenho da ferramenta computacional implementada como parte integrante da metodologia. A análise experimental procurou mensurar a capacidade da solução em coletar dados relevantes de diversas fontes OSINT, categorizá-los segundo as dimensões do 5W3H, e oferecer ao analista de segurança uma visão ampliada, estruturada e operacionalmente útil da ameaça observada.

Os resultados obtidos foram analisados quanto à completude das informações extraídas, à relevância dos dados para a caracterização do IoC, à coerência entre as fontes consultadas e ao nível de detalhamento fornecido em cada dimensão do framework. Dessa forma, a etapa experimental também teve o papel de identificar eventuais limitações da solução proposta, oferecendo subsídios para sua melhoria contínua.

5.2 CENÁRIO E PARÂMETROS DO ESTUDO DE CASO

Para a validação prática do framework proposto e da ferramenta EnricherV2, foi conduzido um estudo de caso centrado na análise de dois endereços IP selecionados como IoCs. Esses IPs foram escolhidos com base em sua reputação negativa em fontes públicas de segurança, estando associados a atividades suspeitas ou maliciosas conforme registros em bancos de dados colaborativos e plataformas de inteligência de ameaças.

O enriquecimento dos IPs foi realizado em quatro cenários distintos, cada um representando uma abordagem diferente para coleta e contextualização das informações:

- Cenário 1: Enriquecimento manual, por meio de consultas diretas às principais fontes públicas e OSINT.

- Cenário 2: Utilização do framework automatizado EnricherV2, desenvolvido neste trabalho.
- Cenário 3: Enriquecimento via OTX AlienVault.
- Cenário 4: Enriquecimento via Pulsedive.

Essa abordagem comparativa permitiu a análise crítica da cobertura, profundidade e operacionalidade de cada solução, com ênfase nos ganhos e limitações do método proposto em relação a soluções amplamente empregadas no mercado.

5.3 MÉTRICAS DE AVALIAÇÃO

Para assegurar a comparabilidade dos resultados, foram definidas as seguintes métricas principais, tomando como referência as dimensões do modelo 5W3H:

- Cobertura 5W3H: Número de dimensões do framework preenchidas com informações relevantes para cada IP, em cada cenário.
- Profundidade e Qualidade da Informação: Nível de detalhamento, consistência e utilidade dos dados obtidos em cada dimensão.
- Usabilidade e Complexidade Operacional: Quantidade de etapas envolvidas, facilidade de execução e suscetibilidade a erros humanos.

Outras métricas complementares, como taxa de falsos positivos/negativos e reprodutibilidade dos resultados, também foram observadas qualitativamente quando pertinente.

5.4 EXECUÇÃO DOS EXPERIMENTOS

O procedimento experimental consistiu na aplicação sequencial de cada abordagem aos mesmos dois IPs, coletando e organizando os resultados de acordo com as dimensões do modelo 5W3H. Todos os dados obtidos foram registrados de forma sistemática e analisados com base nas métricas estabelecidas, permitindo a construção de tabelas comparativas e a elaboração de uma discussão crítica sobre o desempenho relativo de cada solução.

Os endereços IPs analisados foram:

- 102.130.117.167
- 194.213.18.231

5.4.1 CENÁRIO 1 - MANUAL

Neste cenário, o processo de enriquecimento dos IoCs (IPs) foi realizado manualmente, simulando a rotina tradicional de analistas que recorrem a consultas diretas em múltiplas fontes públicas de Threat Intelligence. Foram acessadas bases OSINT relevantes, de modo independente e sem o uso de automação. Para cada IP, as informações obtidas foram coletadas, analisadas e organizadas segundo as dimensões do modelo 5W3H, demandando a atuação ativa do analista em todas as etapas, desde a busca inicial até a consolidação dos dados em relatório estruturado.

5.4.1.1 IP 102.130.117.167

Resultados da coleta manual - Tabela 5.1

Tabela 5.1: Threat Intelligence Summary for IoC (ip) 102.130.117.167

Aspect	Details
What	**Threat:** TOR NODE - Passagem para brute force, bad web bot, etc... Identificado arquivo com tags Trojan, Ransomware, Dropper, Stealer **Files:** trojan.osiris/safebits; 3e4a6cc36d0ddaf95b383ddf704b00c81bac2ac70288f77e307b53ea3b99e561.exe **Hashes:** md5: c2bd4d8ed1b2777cf22abdf6a9bdda29 **TOR Network:** Exit Node **OS:** ...
Who	**Domain:** hostafrica.com **Hostname:** sortie-tor.a-n-o-n-y-m-e.net **ISP:** HA-VPS-NET
Where	**Geolocation:** Latitude: -26.2022; Longitude: 28.0436 **City - Country:** Johannesburg, Gauteng - South Africa. **Associated Address:** ... **Reported in:** SP, SW, US, CH, GE, MK, SG, BE, CZ, CA, FR, HU **Impacted devices:** Windows **Target Countries:** SP, SW, US, CH, GE, MK, SG, BE, CZ, CA, FR, HU
When	**Created:** Arquivo vinculado criado em 19/06/1992 **First submission(file):**28/12/2021 **First reported (IP):** 18/04/2024 **Last reported:** 30/06/2025
Why	**Motivation:** Anonimato, Financeiro
How	**Tools:** Osiris Banking Trojan **TTP [MITRE]:** Execution TA0002; Persistence TA0003; Privilege Escalation TA0004; Defense Evasion TA0005; Credential Access TA0006; Discovery TA0007; Collection TA0009; Command and Control TA0011; Impact TA0040
How Much	**Damage:** Nothing found. **Operational disruptions:** Nothing found.
How Long	**Times reported:** 100

5.4.1.2 IP 194.213.18.231

Resultado da coleta manual - Tabela 5.2

Tabela 5.2: Threat Intelligence Summary for IoC (ip) 194.213.18.231

Aspect	Details
What	**Threat:** botnet cc **Files:** FAKEUPDATES **Hashes:** ... **TOR Network:** NÃO **OS:** -
Who	**Domain:** publynx.com **Hostname:** www.publynx.com **ISP:** Clouvider Limited
Where	**Geolocation:** 39.0395, -77.4918 **City - Country:** Virginia - Ashburn - EUA **Associated Address:** - - **Reported in:** US JP TW BR DE FR GB CA AU ES IT **Impacted devices:** Principalmente Sistemas Windows **Target Countries:** US JP TW BR DE FR GB CA AU ES IT
When	**Created:** ... **First reported:** 06/06/2025 **Last reported:** ...
Why	**Motivation:** Economic motivations.
How	**Tools:** FAKEUPDATES [Downloader - Botnet cc] **TTP [Mitre]:** T1189 – Drive-by Compromise; T1204.001 – User Execution: Malicious Link; T1059 – Command and Scripting Interpreter (JavaScript); T1036.005 – Masquerading: Match Legitimate Resource Name or Location; T1027.013 – Obfuscated Files or Information: Encrypted/Encoded File; T1027.015 – Obfuscated Files or Information: Compression; T1482 – Domain Trust Discovery; T1518 – Software Discovery; T1082 – System Information Discovery; T1033 – System Owner/User Discovery; T1057 – Process Discovery; T1016 – System Network Configuration Discovery; T1614 – System Location Discovery; T1105 – Ingress Tool Transfer; T1074.001 – Data Staged: Local Data Staging; T1048.003 – Exfiltration Over Alternative Protocol: Exfiltration Over Unencrypted Non-C2 Protocol; T1047 – Windows Management Instrumentation; T1102 – Web Service
How Much	**Damage:** Nothing found. **Operational disruptions:** Nothing found.
How Long	**Times reported:** Menos que 100 vezes desde 06/2025

5.4.2 CENÁRIO 2 - ENRICHERV2

Neste segundo cenário, os mesmos IPs foram submetidos ao processo de enriquecimento utilizando a ferramenta automatizada EnricherV2, desenvolvida neste trabalho. A ferramenta integra e correlaciona, de forma automática, dados provenientes de múltiplas fontes OSINT, categorizando-os conforme o modelo 5W3H. O procedimento consistiu na execução do EnricherV2 para cada IP, com registro dos resultados gerados pela ferramenta e posterior validação dos dados extraídos. Esse cenário teve como objetivo avaliar

o ganho de eficiência, padronização e completude proporcionados pela automação em comparação ao método manual.

5.4.2.1 IP 102.130.117.167

Resultados do EnricherV2

- What: Trata-se de IP vinculado a rede TOR. Especificamente neste ponto não é definido o tipo de ameaça.

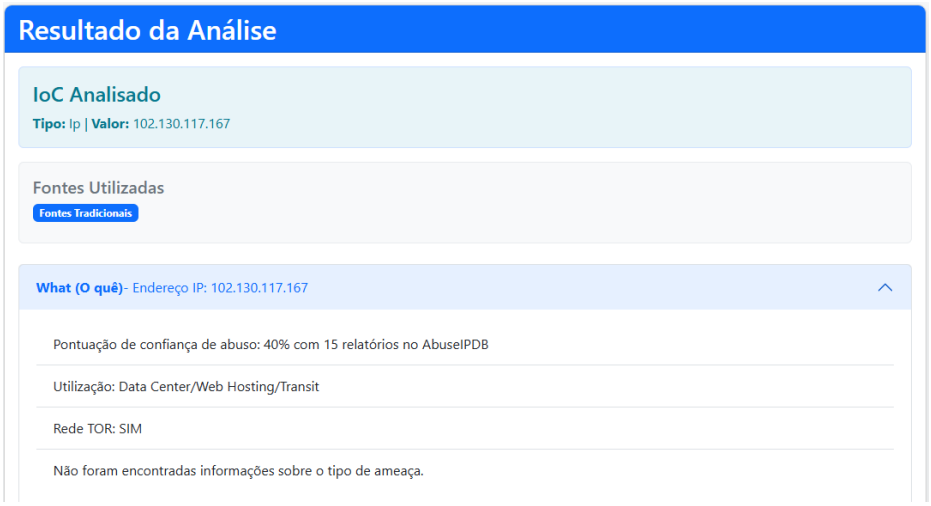

Figura 5.1: Caption

- Who: IP vinculado à Host Africa Ltd

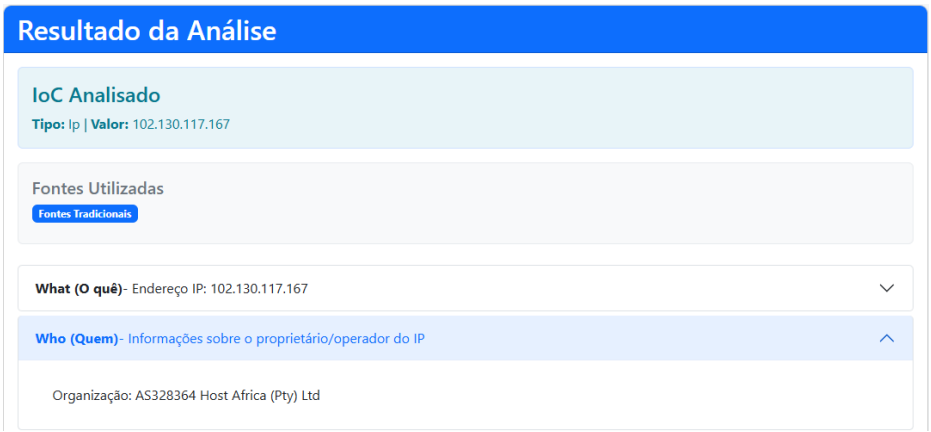

Figura 5.2: Caption

- Where: Localização Geográfica do IP - Johannesburg, Gauteng, África do Sul.

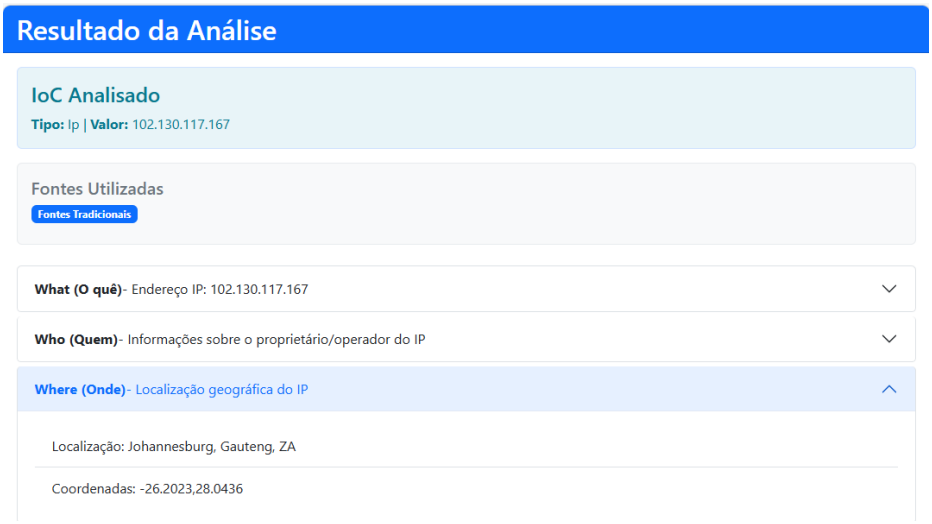

Figura 5.3: Atributo Where

- When: A ferramenta retorna a informação que o IP teve sua última ativiade reportada em 10 de junho de 2025. Frisa-se aqui a necessidade de se reportar o início das atividades.

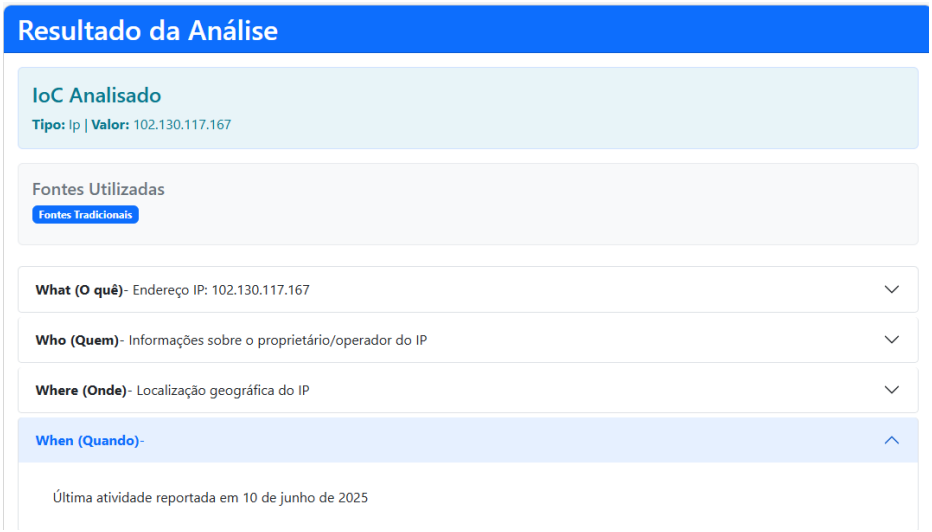

Figura 5.4: Atributo When

- Why: Neste quesito a ferramenta informa o possível uso para anonimização de tráfego e atividades maliciosas.

Figura 5.5: Atributo Why

- How: É informada a utilização como nó da rede TOR para ocultação da origem do tráfego.

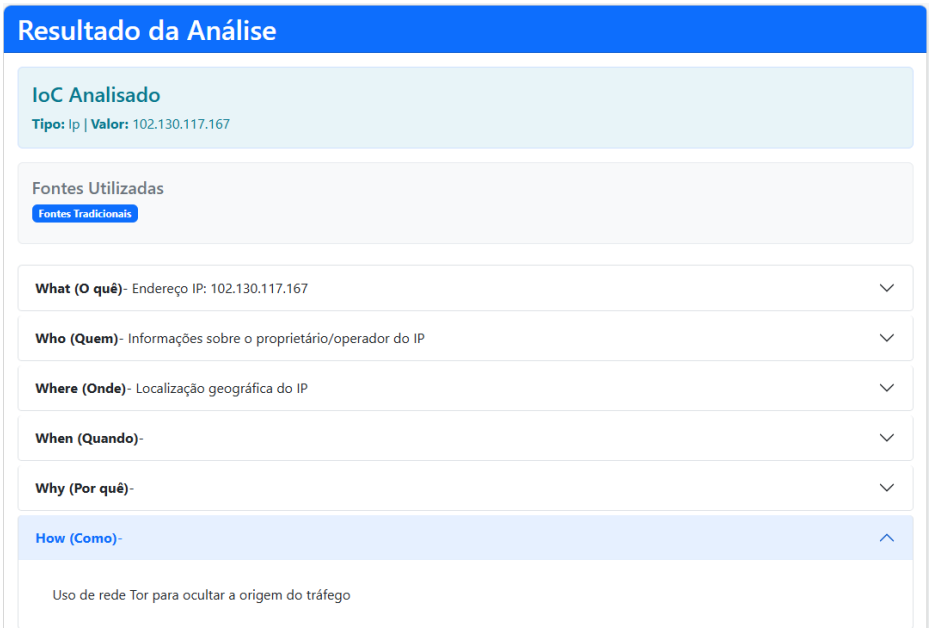

Figura 5.6: Atributo How

- How Much: Neste ponto, a ferramenta considera como impacto o score do AbuseIPDB e seus relatos de abuso.

Figura 5.7: Atributo How much

- How Long: Por fim é retornado pela ferramenta que os relatos permanecem ativos, com reportes de usuários distintos.

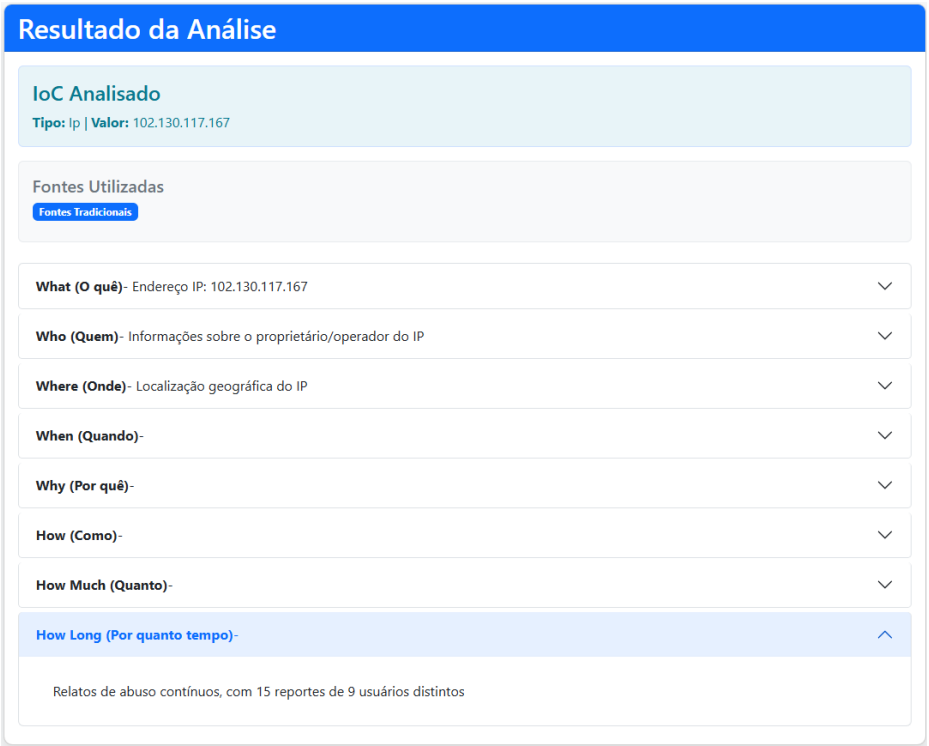

Figura 5.8: Atributo How long

5.4.2.2 IP 194.213.18.231

Resultados do EnricherV2

- What: Trata-se de IP vinculado a servidor de Comando e Controle de botnet, tendo como Malware associado o FAKEUPDATES.

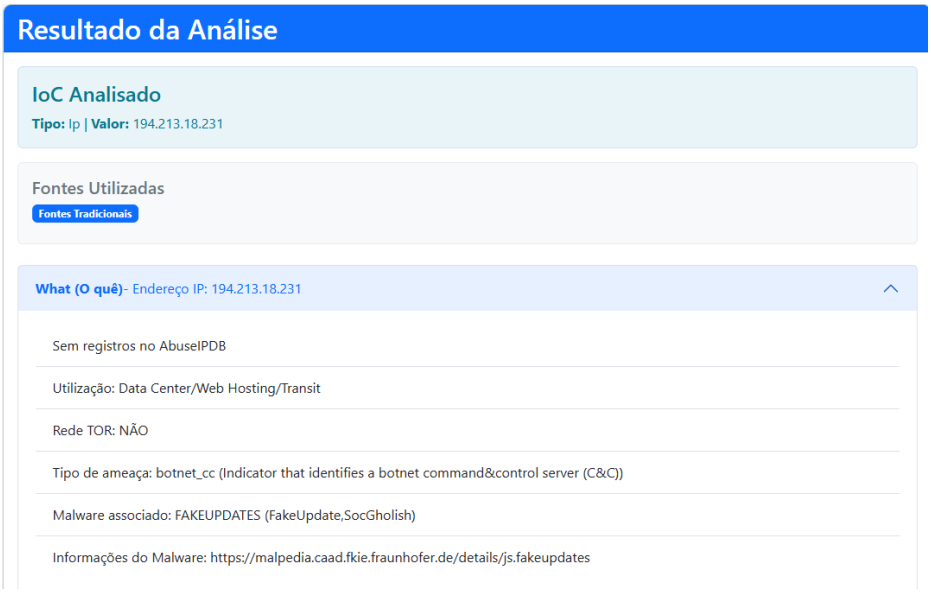

Figura 5.9: Caption

- Who: IP vinculado à Organização AS62240 Clouvider.

Figura 5.10: Caption

- Where: Localização Geográfica do IP - Ashburn, Virginia, US.

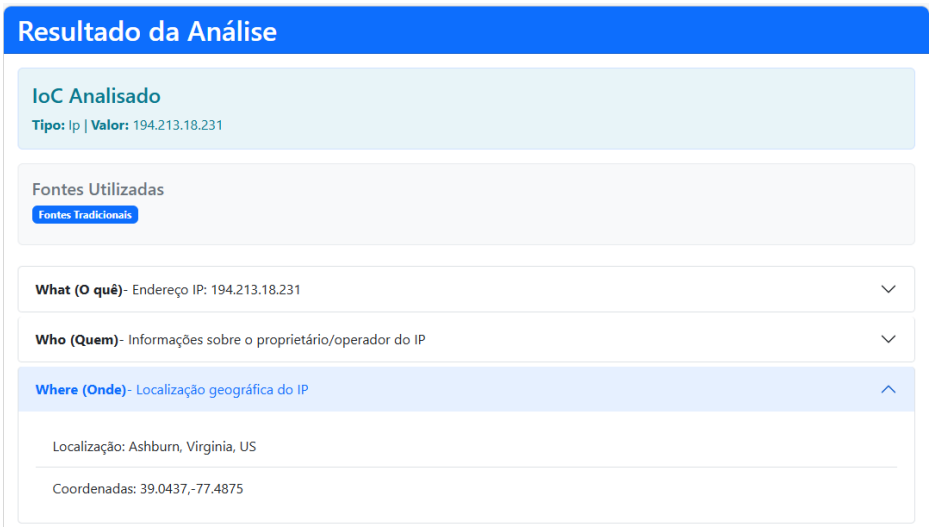

Figura 5.11: Atributo Where

- When: Há registros de atividades detectadas em 27 de setembro de 2022 e novamente em 06 de junho de 2025. Mas não há informação do intervalo entre essas duas datas.

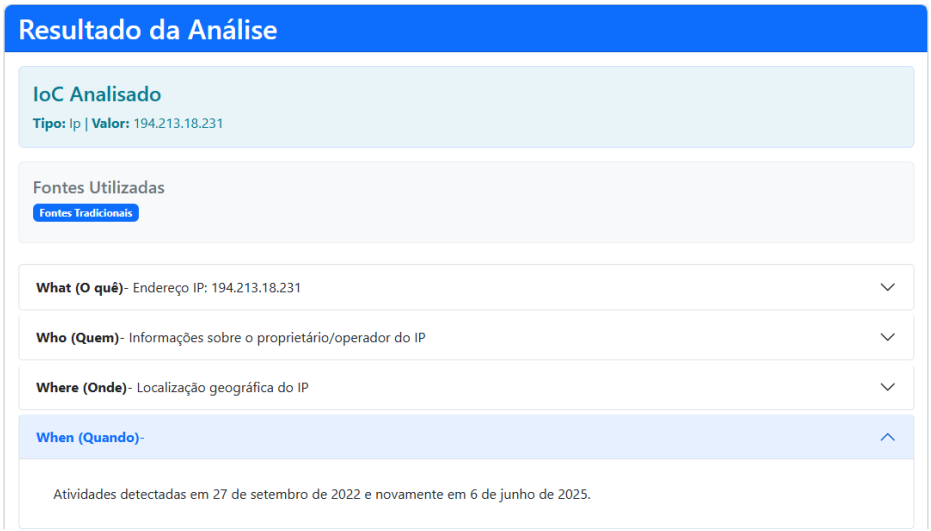

Figura 5.12: Atributo When

- Why: Neste quesito a ferramenta informa o possível uso para controle remoto de dispositivos infectados e distribuição de malware.

Figura 5.13: Atributo Why

- How: É informada, novamente, a utilização como C2 para Botnet, além de exploração de vulnerabilidades (como CVE-2017-17215) e dropper Unix.Mirai. A ferramenta aponta ainda Técnicas associadas ao Mitre ATT&CK: T1105 e T1071.

Figura 5.14: Atributo How

- How Much: Neste ponto, a ferramenta supõe um impacto alto devido ao controle remoto de dispositivos e possível distribuição de malware.

Figura 5.15: Atributo How much

- How Long: Finalmente, é retornado pela ferramenta que a atividade foi detectada por pelo menos 1 dia em junho de 2025, mas com histórico de atividade anterior a 2022.

Figura 5.16: Atributo How long

5.4.3 CENÁRIO 3 - OTX ALIENVAULT

No terceiro cenário, os IPs foram enriquecidos exclusivamente por meio da plataforma OTX AlienVault [otx.alienvault.com], uma das soluções colaborativas mais amplamente empregadas para análise e contextualização de IoCs. O procedimento envolveu a submissão manual dos IPs à interface da plataforma, com coleta dos dados disponibilizados e categorização dos resultados segundo o framework 5W3H. Buscou-se avaliar a cobertura, profundidade e relevância das informações que podem ser obtidas utilizando-se apenas essa fonte, refletindo a experiência de organizações que fazem uso direto do OTX em suas rotinas de Threat Intelligence. As Figuras X1 e X2 apresentam as telas das soluções. Os resultados podem ser vistos nas Tabelas 5.3 e 5.4.

Tabela 5.3: Preenchimento das dimensões 5W3H para o IP 102.130.117.167 na plataforma OTX AlienVault.

Dimensão	**Informação OTX AlienVault**
What	TOR Node
Who	AS328364 african network information center (2 TLDs)
When	First seen: 16/11/2019
Where	Johannesburg, South Africa
Why	...
How	...
How Much	...
How Long	16/11/2019 to 26/01/2024

Tabela 5.4: Preenchimento das dimensões 5W3H para o IP 194.213.18.231 na plataforma OTX AlienVault.

Dimensão	**Informação OTX AlienVault**
What	Unix.Dropper.Mirai-7135870-0 Hash: 81af7c54de544504bfaf417374c9e379c9c57c018ca203b3ae63ec873b013cd1
Who	www.publynx.com | sandboxteam.cyou
When	First seen: 10/06/2023
Where	United Kingdom of Great Britain and Northern Ireland
Why	...
How	...
How Much	...
How Long	10/06/2023 to 06/06/2025

5.4.4 CENÁRIO 4 - PULSEDIVE

No quarto cenário experimental, foi utilizada a plataforma Pulsedive para o enriquecimento dos IPs selecionados. A Pulsedive agrega e classifica informações de múltiplas fontes, oferecendo contexto detalhado sobre cada IoC analisado. O processo consistiu na consulta manual dos IPs na plataforma, seguida da extração e categorização das informações retornadas de acordo com as dimensões do modelo 5W3H. Este cenário permitiu observar o perfil dos dados disponibilizados por essa solução, assim como seu potencial de complementaridade frente às demais abordagens.

Os resultados podem ser vistos nas Tabelas 5.5 e 5.6.

Os resultados obtidos na aplicação do framework foram sistematicamente comparados aos de três ou-

Tabela 5.5: Preenchimento das dimensões 5W3H para o IP 102.130.117.167 na plataforma PulseDive.

Dimensão	Informação PulseDive
What	TOR Node
Who	https://a-n-o-n-y-m-e.net/
When	Added: 17/04/2024
Where	Johannesburg, South Africa
Why	—
How	—
How Much	—
How Long	17/04/2024 to 02/07/2025

Tabela 5.6: Preenchimento das dimensões 5W3H para o IP 194.213.18.231 na plataforma PulseDive.

Dimensão	Informação PulseDive
What	—
Who	—
When	—
Where	Ashburn, Virginia, Clouvider Limited
Why	—
How	—
How Much	—
How Long	—

tros cenários experimentais, abrangendo a ferramenta EnricherV2 e duas outras plataformas amplamente utilizadas para enriquecimento de IoCs. O cenário controle foi constituído pela coleta manual das informações, por meio de consultas individuais aos principais repositórios OSINT, com o apoio do framework. Em sequência, foram avaliados os resultados gerados pela ferramenta EnricherV2 desenvolvida neste trabalho, bem como aqueles provenientes de duas plataformas especializadas: OTX AlienVault e Pulsedive.

A comparação entre os cenários buscou evidenciar as diferenças em termos de eficiência, abrangência, detalhamento e reprodutibilidade das informações obtidas para cada IP analisado. Observou-se, ao longo dos experimentos, variação tanto na quantidade quanto na qualidade dos dados extraídos por cada abordagem, o que permitiu identificar vantagens e limitações específicas de cada método. Os resultados demonstram o panorama atual do processo de enriquecimento de IoCs e servirão de base para a discussão seguinte, onde serão analisadas as principais contribuições, restrições e implicações práticas das soluções avaliadas.

5.5 DISCUSSÃO DOS RESULTADOS

A avaliação dos quatro cenários experimentais evidenciou o impacto do uso do framework proposto frente às abordagens tradicionais e plataformas amplamente adotadas para enriquecimento de CTI. Nos cenários 1 (manual) e 2 (EnricherV2), foi possível identificar todos os parâmetros das dimensões 5W3H para ambos os IPs analisados, demonstrando a capacidade da metodologia em promover uma contextualização completa e estruturada dos indicadores. O processo, tanto manual quanto automatizado via EnricherV2,

Tabela 5.7: Preenchimento das dimensões 5W3H para o IP 1 nas soluções OTX AlienVault, Pulsedive e EnricherV2.

Dimensão 5W3H	OTX AlienVault	Pulsedive	EnricherV2	Manual
What	✓	✓	✓	✓
Who	✓	–	✓	✓
When	✓	–	✓	✓
Where	✓	✓	✓	✓
Why	✓	✓	✓	✓
How	✓	–	✓	✓
How Much	–	–	–	✓
How Long	✓	✓	–	✓

Obs.: "–" indica que a dimensão não foi preenchida pela solução.

permitiu não só reunir, mas também organizar e correlacionar informações dispersas em narrativas analíticas coesas e operacionalmente relevantes.

Em contrapartida, os resultados obtidos nos cenários baseados em plataformas especializadas – OTX AlienVault (cenário 3) e Pulsedive (cenário 4) – evidenciaram limitações importantes. A análise dos IPs nessas plataformas revelou a ausência frequente de informações em parâmetros qualitativos, como "Why", "How" e "How Much", mesmo quando atributos essenciais como "What", "Who", "When" e "Where" estavam razoavelmente preenchidos. Por exemplo, no OTX AlienVault, ambos os IPs analisados apresentaram informações detalhadas sobre o tipo de ameaça, localização e histórico, mas careceram de dados relativos à motivação, técnica empregada ou impacto. No caso da Pulsedive, as lacunas foram ainda mais expressivas, com vários parâmetros não preenchidos, especialmente para o segundo IP.

Essas observações reforçam que, embora as plataformas comerciais e colaborativas de Threat Intelligence desempenhem papel fundamental no ecossistema de CTI, elas tendem a focar nos aspectos técnicos e descritivos dos IoCs, não conseguindo, muitas vezes, cobrir todas as dimensões necessárias para uma análise aprofundada e estratégica. Nesse sentido, a aplicação do framework proposto mostrou-se decisiva para ampliar a completude dos relatórios, aumentar o grau de contextualização e proporcionar subsídios mais robustos para a tomada de decisão.

Entre os principais avanços demonstrados estão a capacidade do framework de estruturar os dados em múltiplas dimensões, facilitar a automação da coleta e organização, e permitir a adaptação a diferentes perfis de IoCs. A integração de diversas fontes OSINT, com priorização do preenchimento de "What", "Who" e "Where", seguida de iteração nas demais dimensões, mostrou-se um caminho eficiente para mitigar as limitações inerentes à dispersão ou ausência de dados.

Por outro lado, os experimentos também confirmaram desafios já reconhecidos na literatura, como a dependência de fontes públicas, a necessidade de intervenção analítica para interpretação de parâmetros subjetivos ("Why", "How Much") e as limitações quanto à automação plena em ambientes dinâmicos ou de grande escala. Tais aspectos apontam para a necessidade de avanços futuros, como a incorporação de algoritmos de machine learning, integração com fontes privadas e uso de modelos avançados para inferência e sumarização de informações.

Em síntese, os resultados obtidos ressaltam a relevância de abordagens estruturadas e multidimensio-

nais no enriquecimento de CTI e apontam caminhos para o desenvolvimento de soluções mais completas, escaláveis e alinhadas às demandas reais dos profissionais de segurança da informação.

5.6 DISCUSSÃO COMPARATIVA COM A LITERATURA

Para avaliar o posicionamento e as contribuições do framework proposto frente ao estado da arte, realizou-se uma análise comparativa com os principais trabalhos correlatos apresentados na Seção 2. As Tabelas 2.1 e 2.2 já resumem os principais métodos, escopo, limitações e avanços de cada proposta revisada. No entanto, é fundamental retomar essa discussão à luz dos resultados experimentais obtidos nesta pesquisa.

A Tabela 5.8 sintetiza os principais aspectos das soluções analisadas, comparando-as em termos de automação, cobertura de fontes, contextualização, integração, uso de estrutura multidimensional (como o 5W3H), avaliação de resultados e aplicabilidade prática.

Tabela 5.8: Comparação entre o framework proposto e trabalhos correlatos selecionados.

Proposta	**Automação**	**Multi-Fonte**	**Contexto (5W3H)**	**Avaliação**	**Integr.**	**Diferencial**
Este trabalho	Parcial	Sim	Completa	Experimental	Sim	Estrutura 5W3H, customização
TIMiner (38)	Alta	Social	Moderada	Experimental	Parcial	Extração automatizada, ML
ETIP (16)	Alta	Sim	Score/contexto	Experimental	Alta	Integração SIEM, scoring
Osliak (13)	Média	Sim	Políticas dinâmicas	Experimental	Alta	Foco ICS/SCADA
Park (43)	Modular	Sim	Metadados	Experimental	Sim	Extensibilidade IoT
Slinde (14)	Baixa	Sim	Ciclo Intel.	Qualitativa	Não	Foco processo e requisitos
Sun (15)	Variável	Sim	Panorama	Survey	Parcial	Revisão/Desafios atuais

Os resultados experimentais demonstraram que, ao contrário das plataformas automatizadas como OTX AlienVault e Pulsedive, o framework proposto foi capaz de preencher de forma consistente todas as dimensões do modelo 5W3H para os casos avaliados, agregando valor ao processo de contextualização e análise de IoCs. Essa completude só foi possível graças à integração de múltiplas fontes OSINT e à organização estruturada dos dados, diferindo de métodos que se concentram apenas na extração automatizada de indicadores ou em abordagens generalistas de classificação.

Comparativamente, soluções como o TIMiner (38) e ETIP (16) destacam-se pelo alto grau de automação e integração com sistemas SIEM, além de implementar mecanismos de scoring automático para priorização de ameaças. No entanto, tais soluções tendem a apresentar limitações quanto à contextualização detalhada, principalmente em parâmetros subjetivos como motivações (“Why”) e impacto (“How Much”), lacunas igualmente observadas nas plataformas testadas nesta pesquisa.

A abordagem adotada por Osliak (13), ao focar em infraestruturas críticas, contribui com políticas dinâmicas baseadas em CTI, porém restringe sua aplicabilidade a cenários muito específicos. Park (43) propõe um esquema extensível e modular para IoT, o que pode inspirar futuras adaptações do presente framework para contextos emergentes. Por sua vez, Slinde (14) enfatiza a importância do ciclo de inteligência e de requisitos claros de OSINT, aspectos contemplados nesta pesquisa por meio do uso do ciclo de inteligência e da metodologia 5W3H.

Por fim, a revisão de Sun (15) ressalta os desafios persistentes na mineração e aplicação prática de CTI, como a sobrecarga de dados irrelevantes, a dificuldade de automação e a validação de fontes. O framework proposto neste trabalho busca mitigar esses desafios ao estruturar o fluxo de coleta, verificação e categorização das informações, promovendo uma inteligência acionável e adaptável às necessidades organizacionais.

Em síntese, a comparação evidencia que, embora existam soluções avançadas e automatizadas no campo do enriquecimento de CTI, o diferencial do framework proposto reside na sua capacidade de contextualização multidimensional, adaptabilidade e potencial de integração prática a diferentes ambientes de cibersegurança. Esse posicionamento reforça a relevância e a originalidade do método frente ao panorama da literatura atual, ao mesmo tempo em que aponta caminhos claros para aprimoramentos futuros, como maior automação, integração de mecanismos de scoring e adaptação a novos domínios tecnológicos.

6 CONCLUSÃO

O presente trabalho propôs, desenvolveu e avaliou um framework estruturado para o enriquecimento de CTI utilizando dados provenientes de OSINT. Fundamentado no modelo 5W3H e apoiado pelo desenvolvimento de uma ferramenta computacional própria (EnricherV2), o estudo buscou avançar no estado da arte ao oferecer uma abordagem sistemática, replicável e eficiente para a integração, organização e análise de múltiplas dimensões de informações sobre Indicadores de Comprometimento (IoCs).

Os experimentos realizados demonstraram que, tanto por meio do método manual quanto com a ferramenta automatizada, foi possível preencher todas as dimensões do framework 5W3H para os casos testados. Essa completude permitiu construir relatórios analíticos mais robustos, trazendo contexto operacional, histórico, infraestrutura envolvida, motivações e possíveis impactos associados a cada IoC, superando as limitações observadas em abordagens tradicionais ou centradas apenas em plataformas automatizadas. A aplicação prática evidenciou, ainda, a relevância de uma metodologia estruturada não apenas para enriquecer tecnicamente os dados, mas também para apoiar decisões estratégicas em cibersegurança e resposta a incidentes.

A comparação com plataformas especializadas de mercado, como OTX AlienVault e Pulsedive, ressaltou os diferenciais da abordagem proposta. Embora essas soluções se destaquem pela praticidade e cobertura técnica, frequentemente apresentaram lacunas em dimensões qualitativas fundamentais para a análise contextual, como motivações (“Why”), métodos (“How”) e magnitude dos impactos (“How Much”, How Long"). Tais limitações evidenciam a importância de frameworks que priorizem a integração inteligente e a análise multidimensional, para além da simples agregação de dados de reputação.

Não obstante os avanços, o estudo reconheceu desafios inerentes ao uso de fontes abertas, como a variação na qualidade e atualização dos dados, as restrições de acesso e o potencial de viés em informações públicas. Além disso, a necessidade de intervenção analítica para interpretar e validar dimensões subjetivas permanece um ponto sensível, mesmo com o suporte de automação. Esses aspectos apontam para oportunidades futuras, como a integração de algoritmos de aprendizado de máquina, expansão de fontes (inclusive privadas e comerciais), e o uso de modelos de linguagem para inferência contextual e sumarização.

Como contribuições principais, este trabalho entrega uma metodologia testada e uma ferramenta desenvolvida, capazes de elevar o padrão de enriquecimento em CTI via OSINT. Tais resultados são de grande utilidade tanto para a comunidade acadêmica quanto para equipes operacionais, promovendo maior resiliência cibernética e subsidiando políticas de defesa mais informadas e eficientes.

Recomenda-se, como próximos passos, a ampliação do estudo para outros tipos de IoCs, a avaliação do framework em cenários dinâmicos e de larga escala, e o aprofundamento da integração com tecnologias emergentes em inteligência artificial e computação quântica. Por fim, destaca-se que o sucesso das estratégias de cibersegurança dependerá, cada vez mais, da capacidade de transformar grandes volumes de dados abertos em inteligência contextualizada, prática e acionável, papel para o qual a metodologia proposta demonstrou-se promissora.

REFERÊNCIAS BIBLIOGRÁFICAS

1 TANABE, R.; OLIVEIRA-ALBUQUERQUE, R. de; SILVA-FILHO, D. da; SILVA, D. Alves-da; COSTA-GONDIM, J.-J. Osint methods in the intelligence cycle. In: GARCIA, M. V.; GORDÓN-GALLEGOS, C. (Ed.). *CSEI: International Conference on Computer Science, Electronics and Industrial Engineering (CSEI)*. Cham: Springer Nature Switzerland, 2023. p. 42–54. ISBN 978-3-031-30592-4.

2 BIANCO, D. *The Pyramid of Pain*. 2013. Blog post. Disponível em: <https://detect-respond.blogspot.com/2013/03/the-pyramid-of-pain.html>.

3 Schlette, D.; Caselli, M.; Pernul, G. A comparative study on cyber threat intelligence: The security incident response perspective. *IEEE Communications Surveys and Tutorials*, v. 23, n. 4, p. 2525–2556, 2021.

4 Verizon Business. *2024 Data Breach Investigations Report – Summary of Findings*. 2024. <https://www.verizon.com/business/en-nl/resources/reports/dbir/2024/summary-of-findings/>. Accessed: 2025-08-11.

5 Verizon Business. *Vulnerability Exploitation Boom – Supply Chain Breaches Up 68%*. 2024. <https://www.verizon.com/about/news/2024-data-breach-investigations-report-vulnerability-exploitation-boom>. Accessed: 2025-08-11.

6 IBM Security X-Force. *X-Force Threat Intelligence Index 2024*. 2024. <https://www.ibm.com/think/x-force/2024-x-force-threat-intelligence-index>. Accessed: 2025-08-11.

7 International Data Corporation. *Worldwide Digital Transformation Spending Guide*. 2024. <https://www.businesswire.com/news/home/20231101754700/en/Worldwide-Digital-Transformation-Spending-Forecast-to-Continue-Its-Double-Digit-Growth-Trajectory-According-to-
Accessed: 2025-08-11.

8 Security Agency, I. *The President's National Security Telecommunications Advisory Committee Draft NSTAC Report to the President Measuring and Incentivizing the Adoption of Cybersecurity Best Practices*. 2021.

9 Ghioni, R.; Taddeo, M.; Floridi, L. Open source intelligence and ai: a systematic review of the gelsi literature. *AI and Society*, 2023.

10 EXéRCITO BRASILEIRO, COMANDO DE OPERAçõES TERRESTRES. *Manual de Campanha MC 2.40-54: Inteligência de Fontes Abertas*. 1ª edição. ed. Brasília, 2025. Aprovado pela Portaria – COTER/C Ex Nº 531, de 24 de abril de 2025.

11 PUYVELDE, D. V.; RIENZI, F. T. The rise of open-source intelligence. *European Journal of International Security*, 2025. Advance online publication. Disponível em: <https://www.cambridge.org/core/journals/european-journal-of-international-security/article/rise-of-open-source-intelligence/8F4E6BDC05D58A967727841482BF3AC8>.

12 Yadav, A.; Kumar, A.; Singh, V. Open-source intelligence: a comprehensive review of the current state, applications and future perspectives in cyber security. *Artificial Intelligence Review*, v. 56, n. 11, p. 12407–12438, 2023.

13 Osliak, O.; Saracino, A.; Martinelli, F.; Mori, P. Cyber threat intelligence for critical infrastructure security. *Concurrency and Computation: Practice and Experience*, v. 35, n. 23, 2023.

14 Sofie, J.; Supervisor, S.; Radianti, J. *Unveiling the Potential of Open-Source Intelligence (OSINT) for Enhanced Cybersecurity Posture: A study on OSINT implementation and utilization within organizations and recommendations for increased leverage of OSINT's advantages*. 2021.

15 Sun, N.; Zhang, Y.; Wu, X.; Wang, M.; Yu, L.; Chen, B. Cyber threat intelligence mining for proactive cybersecurity defense: A survey and new perspectives. *IEEE Communications Surveys and Tutorials*, v. 25, n. 3, p. 1748–1774, 2023.

16 González-Granadillo, G.; Faiella, M.; Medeiros, I.; Azevedo, R.; González-Zarzosa, S. Etip: An enriched threat intelligence platform for improving osint correlation, analysis, visualization and sharing capabilities. *Journal of Information Security and Applications*, v. 58, May 2021.

17 Tatam, M.; Shanmugam, B.; Azam, S.; Kannoorpatti, K. *A review of threat modelling approaches for APT-style attacks*. 2021.

18 da Silva, R. M.; Gondim, J. J. C.; de O. Albuquerque, R. *Methodology to Improve the Quality of Cyber Threat Intelligence Production Through Open Source Platforms*. 2021.

19 SAEED, S.; SUAYYID, S. A.; AL-GHAMDI, M. S.; AL-MUHAISEN, H.; ALMUHAIDEB, A. M. A systematic literature review on cyber threat intelligence for organizational cybersecurity resilience. *Sensors*, v. 23, n. 16, 2023. ISSN 1424-8220. Disponível em: <https://www.mdpi.com/1424-8220/23/16/7273>.

20 THE NIST Cybersecurity Framework (CSF) 2.0. [S.l.], 2024. Final version released Feb 26, 2024. Disponível em: <https://nvlpubs.nist.gov/nistpubs/CSWP/NIST.CSWP.29.pdf>.

21 INTERNATIONAL ORGANIZATION FOR STANDARDIZATION (ISO) / INTERNATIONAL ELECTROTECHNICAL COMMISSION (IEC). *Information security, cybersecurity and privacy protection — Information security management systems — Requirements*. Geneva, 2022. Annex A includes control 5.7 "Threat intelligence" aligned to ISO/IEC 27002:2022.

22 BROWN, R.; NICKELS, K. *2023 SANS Cyber Threat Intelligence (CTI) Survey: Keeping Up with a Changing Threat Landscape*. 2023. SANS White Paper. Disponível em: <https://www.sans.org/white-papers/2023-cti-survey-keeping-up-changing-threat-landscape>.

23 ENISA Threat Landscape 2024. [S.l.], 2024. Disponível em: <https://www.enisa.europa.eu/sites/default/files/2024-11/ENISA%20Threat%20Landscape%202024_0.pdf>.

24 SOUSA, C. E.; ALBUQUERQUE, R. d. O.; GONDIM, J. J. C. Enriched cyber threat intelligence through osint: A methodology for strengthening cybersecurity resilience. In: *Emerging Trends in Information Systems and Technologies, WorldCIST 2025*. Porto, Portugal: Springer, 2025.

25 ALBUQUERQUE, R. d. O.; GONDIM, J. J. C.; SOUSA, C. E. *Enricher*. 2024. Certificado de Registro de Programa de Computador BR512025001037-0. Instituto Nacional da Propriedade Industrial (INPI), Brasil. Programa desenvolvido em Python. Válido por 50 anos a partir de 1º de janeiro subsequente à data de publicação. Disponível em: <https://www.gov.br/inpi/pt-br>.

26 WARNER, M. Wanted: A definition of "intelligence": Understanding our craft. *Studies in Intelligence*, Center for the Study of Intelligence, CIA, v. 46, n. 3, 2002. Disponível em: <https://www.cia.gov/resources/csi/static/Wanted-Definition-of-Intel.pdf>.

27 KENT, S. *Strategic Intelligence for American World Policy*. Princeton University Press, 1966. ISBN 9780691624044. Disponível em: <http://www.jstor.org/stable/j.ctt183q0qt>.

28 INTELIGêNCIA, A. B. de. *Doutrina da Atividade de Inteligência*. 2023. Documento oficial. Disponível em: <https://www.gov.br/abin/pt-br/centrais-de-conteudo/publicacoes/doutrina/Doutrina-da-Atividade-de-Inteligencia-2023> Acesso em: 16 jun. 2025.

29 PINCOVSCY, J.-A.; COSTA-GONDIM, J.-J. Methodology for cyber threat intelligence with sensor integration. In: GARCIA, M. V.; GORDON-GALLEGOS, C. (Ed.). *CSEI: International Conference on Computer Science, Electronics and Industrial Engineering (CSEI)*. Cham: Springer Nature Switzerland, 2023. p. 14–28.

30 DEFENCE, U. M. of. *Joint Doctrine Publication 2-00: Intelligence, Counter-intelligence and Security Support to Joint Operations*. 2023. Joint Doctrine Publication. Disponível em: <https://www.gov.uk/mod/dcdc>. Acesso em: 16 jun. 2025.

31 FORCE, U. S. S. *Space Doctrine Publication 2-0, Intelligence*. 2023. Space Doctrine Publication. OPR: STARCOM Delta 10. Disponível em: <https://www.starcom.spaceforce.mil/Portals/2/Documents/Doctrine/SDP%202-0%20Intelligence.pdf>. Acesso em: 16 jun. 2025.

32 SILVA, R. M. da; GONDIM, J. J. C.; ALBUQUERQUE, R. de O. Methodology to improve the quality of cyber threat intelligence production through open source platforms. In: *Proceedings of the International Conference on Cyber Security (ICCS)*. [S.l.: s.n.], 2022.

33 MACêDO, A.; PEOTTA, L.; GOMES, F. A review of the intersection techniques on humint and osint. *International Journal on Cybernetics & Informatics*, v. 12, n. 1, p. 53–63, 2023.

34 CHAUDHARY, M.; BANSAL, D. Open source intelligence extraction for terrorism-related information: A review. *WIREs Data Mining and Knowledge Discovery*, Wiley, v. 12, n. 5, p. e1473, 2022. Disponível em: <https://doi.org/10.1002/widm.1473>.

35 RAJAMäKI, J.; MCMENAMIN, S. Utilization and sharing of cyber threat intelligence produced by open-source intelligence. *International Conference on Cyber Warfare and Security*, v. 19, p. 607–611, 03 2024.

36 de Melo e Silva, A.; Gondim, J. J. C.; de Oliveira Albuquerque, R.; Villalba, L. J. G. A methodology to evaluate standards and platforms within cyber threat intelligence. *Future Internet*, v. 12, n. 6, June 2020.

37 SATVAT, K.; GJOMEMO, R.; VENKATAKRISHNAN, V. Tipce: A longitudinal threat intelligence platform comprehensiveness analysis. In: *Proceedings of the Fourteenth ACM Conference on Data and Application Security and Privacy*. New York, NY, USA: Association for Computing Machinery, 2024. (CODASPY '24), p. 349–360. ISBN 9798400704215. Disponível em: <https://doi.org/10.1145/3626232.3653278>.

38 Zhao, J.; Yan, Q.; Li, J.; Shao, M.; He, Z.; Li, B. Timiner: Automatically extracting and analyzing categorized cyber threat intelligence from social data. *Computers & Security*, v. 95, August 2020.

39 Hagen, R. A.; Helkala, K. *Complexity of Contemporary Indicators of Compromise*. 2024.

40 A Systems Approach to Indicators of Compromise Utilizing Graph Theory. In: 2018 IEEE International Symposium on Technologies for Homeland Security (HST). [S.l.: s.n.], 2018. p. 1–6.

41 MORIOT, C.; LESUEUR, F.; STOULS, N.; VALOIS, F. How to build socio-organizational information from remote ip addresses to enrich security analysis? In: *2022 IEEE 47th Conference on Local Computer Networks (LCN)*. [S.l.: s.n.], 2022. p. 287–290.

42 SáNCHEZ, L.; LANZA, J.; SANTANA, J. R.; SOTRES, P.; GONZáLEZ, V.; MARTíN, L.; SOLMAZ, G.; KOVACS, E.; DIETZEL, M.; SUMMA, A.; JAFARI, A. R.; MINERVA, R.; CRESPI, N. Data enrichment toolchain: A data linking and enrichment platform for heterogeneous data. *IEEE Access*, v. 11, p. 103079–103091, 2023.

43 Park, Y.; Choi, J.; Choi, J. An extensible data enrichment scheme for providing intelligent services in internet of things environments. *Mobile Information Systems*, 2021.

44 SPYROS, A.; PAPOUTSIS, A.; KORITSAS, I.; MENGIDIS, N.; ILIOU, C.; KAVALLIEROS, D.; TSIKRIKA, T.; VROCHIDIS, S.; KOMPATSIARIS, I. Towards continuous enrichment of cyber threat intelligence: A study on a honeypot dataset. In: *2022 IEEE International Conference on Cyber Security and Resilience (CSR)*. [S.l.: s.n.], 2022. p. 267–272.

45 ARAZZI, M.; R. Arikkat, D.; NICOLAZZO, S.; NOCERA, A.; Rehiman K.A., R.; P., V.; CONTI, M. Nlp-based techniques for cyber threat intelligence. *Computer Science Review*, v. 58, p. 100765, 2025. ISSN 1574-0137. Disponível em: <https://www.sciencedirect.com/science/article/pii/S1574013725000413>.

46 HEVNER, A. R.; MARCH, S. T.; PARK, J.; RAM, S. Design science in information systems research. *MIS Quarterly*, MIS Quarterly, v. 28, n. 1, p. 75–105, 2004. Disponível em: <https://www.jstor.org/stable/25148625>.