IA GENERATIVA APLICADA À AUTOMAÇÃO DE RELATÓRIOS DE CONFORMIDADE EM SEGURANÇA CIBERNÉTICA

Ramicés dos Santos Silva†
PPGCA
Escola Politécnica/UNIVALI
Florianópolis, SC, Brasil
ramices@univali.br

Anita M. da Rocha Fernandes
PPGCA
Escola Politécnica/UNIVALI
Florianópolis, SC, Brasil
anita.fernandes@univali.br

Rudimar L. Scaranto Dazzi
PPGCA
Escola Politécnica/UNIVALI
Itajaí, SC, Brasil
rudimar@univali.br

ABSTRACT

This work addresses the growing demand for efficient solutions in the information security and data protection compliance process, particularly in medium-sized companies that face significant challenges related to the shortage of specialized professionals. This paper implements a solution based on the use of Generative Artificial Intelligence to automate the generation of compliance reports, customized according to the needs of different stakeholders within the organization, such as managers, technical teams, and compliance officers. The main objective is to optimize the efficiency, scalability, and reliability of compliance processes, reducing the time and resources required to produce these documents, in addition to improving the clarity of the presented information. The study's rationale lies in the need to mitigate two main problems faced by companies: the lack of qualified professionals in the cybersecurity field and the workload associated with the manual production of complex reports. These factors limit organizations' ability to implement robust compliance programs and hinder meeting current regulatory requirements. The adopted methodology included a systematic literature review to identify existing applications of Generative AI technologies in the field of information security and data protection, as well as an analysis of best practices in compliance. To validate the proposal, modular architecture was developed, consisting of three main modules: integration of data from security information and event management systems; processing of this data to extract relevant information; and report generation using advanced Generative AI techniques. The results indicate that automating report creation can free up experts to focus on more strategic activities, while improving the accuracy and consistency of the information provided. Furthermore, the results demonstrated that the solution enables the generation of consistent reports, allowing companies to adapt more easily to changes in legal and normative requirements, ensuring greater regulatory compliance and reducing operational risks. Through this approach, implementation represents a significant contribution to the use of artificial intelligence in the business context, promoting a more accessible and sustainable security culture in line with the spirit of Sustainable Development Goal 11 (Sustainable Cities and Communities).

KEYWORDS

Generative Artificial Intelligence, Automation of Reports, Information Security Compliance

1 Introdução

A crescente dependência de sistemas de informação nas organizações tem sido acompanhada por uma intensificação dos riscos cibernéticos, com aumento tanto do volume quanto da sofisticação de ataques e crimes digitais [3]. Em muitos ambientes corporativos, as defesas de segurança não evoluem no mesmo ritmo das técnicas de ataque, em parte porque uma parcela significativa dos processos de segurança da informação ainda depende de atividades manuais realizadas por especialistas.

Diversos estudos apontam para uma escassez global de profissionais qualificados em cibersegurança. Singh et al. [20] descrevem um cenário de déficit persistente de mão de obra especializada, enquanto a Harvard Business Review [5] estima uma lacuna de milhões de profissionais na área. Pesquisas conduzidas por ISC2 [7] sugerem que a falta de tempo e a sobrecarga de tarefas estão entre os fatores mais críticos de insatisfação nas equipes de segurança, e relatórios corporativos indicam que a responsabilidade percebida diante de clientes, reguladores e equipes internas agrava o estresse desses profissionais [6].

Esse quadro é particularmente sensível em empresas de médio porte. Essas organizações operam infraestruturas de TI(Tecnologia da Informação) complexas, precisam demonstrar conformidade com normas de segurança da informação e legislações de proteção de dados e, ao mesmo tempo, contam com equipes enxutas que acumulam funções técnicas, operacionais e de compliance. Do ponto de vista de governança, estruturas típicas envolvem conselho ou sócios, diretoria executiva, jurídico, compliance, auditorias interna e externa e áreas de negócio, exigindo fluxos de informação estruturados e compreensíveis para múltiplos públicos [18].

A produção de relatórios de compliance em segurança da informação e proteção de dados emerge, nesse contexto, como um gargalo relevante. Elaborar esses documentos implica interpretar saídas técnicas de ferramentas como SIEMs, scanners de vulnerabilidades e sistemas de monitoramento, mapear essas

evidências a *frameworks* e normas (como ISO 27001, NIST *Cybersecurity Framework*, COBIT (*Control Objectives for Information and Related Technologies*), GDPR (*General Data Protection Regulation*) e LGPD (Lei Geral de Proteção de Dados)) e traduzir os achados para linguagens adequadas a executivos, profissionais de compliance e equipes técnicas [2, 15]. Esse processo consome tempo expressivo de profissionais já escassos e aumenta a probabilidade de inconsistências, lacunas e divergências entre diferentes relatórios.

Paralelamente, o avanço da Inteligência Artificial Generativa, especialmente na forma de grandes modelos de linguagem (LLMs) baseados em arquiteturas *Transformer* [23], abriu novas possibilidades para automatizar tarefas baseadas em linguagem natural. No plano industrial, reportagens especializadas relatam o uso de modelos generativos em produtos como o Microsoft Security Copilot e o Google Cloud Security AI Workbench, que combinam LLMs com dados de segurança para apoiar analistas na investigação de incidentes e na explicação de alertas [22]. Na academia, trabalhos sobre defesa cibernética autônoma exploram técnicas de aprendizado de máquina, incluindo redes neurais, para detecção e resposta automática a ataques [14]. De forma mais ampla, George, George e Martin [4] discutem o impacto de ferramentas como o ChatGPT sobre o trabalho humano, argumentando que tais ferramentas tendem a automatizar tarefas repetitivas e estruturadas, liberando profissionais para atividades de maior valor agregado.

Diante deste cenário, este artigo propõe o uso de IA (Inteligência Artificial) Generativa como ferramenta de apoio à automação de relatórios de compliance em segurança da informação e proteção de dados, partindo de dados produzidos por plataformas de segurança já existentes, mas tradicionalmente consumidos em linguagem técnica. As questões de pesquisa formuladas investigam como a automação baseada em IA pode contribuir para que empresas de médio porte desenvolvam programas mais robustos de segurança e proteção de dados, se essa abordagem aumenta a escalabilidade da atuação de profissionais de segurança e se os resultados produzidos por IA são suficientemente confiáveis para usos de compliance.

A hipótese central é que uma solução baseada em IA Generativa para geração de relatórios de compliance pode otimizar eficiência, escalabilidade e confiabilidade dessas atividades, desde que seja inserida em uma arquitetura adequada e acompanhada de mecanismos de supervisão humana. Destaca-se ainda que, o trabalho se relaciona ao Objetivo de Desenvolvimento Sustentável 11 da ONU ao fortalecer a capacidade de organizações de médio porte de manter práticas de segurança e proteção de dados mais consistentes, contribuindo para infraestruturas digitais mais seguras e resilientes.

2 Trabalhos Correlatos

Foi realizado um levantamento dos trabalhos relativos ao uso de IA Generativa (GAI/LLMs) em cibersegurança e automação de compliance. Este levantamento é apresentado a seguir.

2.1 Metodologia da revisão da literatura

A revisão da literatura foi conduzida com base no PRISMA (*Preferred Reporting Items for Systematic Reviews and Meta-Analyses*), assegurando transparência, replicabilidade e rigor [17]. O processo contemplou: (i) pesquisa bibliográfica e RSL (Revisão Sistemática da Literatura); (ii) análise de conteúdo e classificação temática; e (iii) síntese e discussão dos resultados. Foram consultadas quatro bases amplamente reconhecidas: Google Scholar[1], IEEE Xplore[2], CAPES Periódicos[3] e Scopus[4]. Para capturar os principais conceitos, utilizaram-se termos de busca como *"Generative AI", "Compliance", "Cyber Security" e "Automatic Report Generation"*, além de uma *string de pesquisa* estruturada: *("generative AI" OR "generative artificial intelligence" OR "LLM" OR "large language model" OR "GPT" OR "ChatGPT" OR "AI generation") AND ("compliance report" OR "automated reporting" OR "report generation" OR "compliance automation" OR "automated compliance") AND ("XDR" OR "extended detection and response" OR "SIEM" OR "security platform" OR "threat detection" OR "security analytics" OR "SOC" OR "security operations")*. As buscas foram realizadas entre setembro e outubro de 2025, limitadas a publicações de 2022 a 2025.

A triagem considerou títulos e resumos para excluir estudos desalinhados aos objetivos, priorizando estudos empíricos e RSLs com aplicações práticas de IA Generativa em cibersegurança; foram excluídos trabalhos exclusivamente sobre regulação de IA, pré-impressões sem revisão por pares e estudos puramente teóricos sem aplicação prática. Na etapa seguinte, procedeu-se à análise textual para avaliar relevância quanto à integração com XDR (*Extended Detection na Response*)/SIEM, implementação de relatórios de conformidade automatizados, metodologia e evidências empíricas. Os critérios de avaliação abrangeram: métricas e especificidade para foco em conformidade e estado da arte em cibersegurança, desempenho/viés, aplicabilidade real, especificidade GAI versus IA/ML genérica e aplicabilidade a empresas de médio porte. Ao final, oito estudos (2023–2025) foram incluídos, com a evolução do funil e distribuição por temas ilustradas na Tabela 1.

Tabela 1. Processo de seleção dos artigos

Base de dados	1ª seleção	2ª seleção	Seleção final
Google Scholar	89	4	2
IEEE Xplore	131	8	3
CAPES	221	12	2
Scopus	69	4	1
Duplicatas removidas	18	0	0
Registros inelegíveis	303	161	20
Total	189	28	8

[1] https://scholar.google.com/
[2] https://ieeexplore.ieee.org/
[3] https://periodicos.capes.gov.br/
[4] https://www.scopus.com/search

2.2 Aplicações de IA Generativa em operações de cibersegurança

Na geração de inteligência de ameaças e relatórios automatizados, Perrina et al. [16] apresentam o AGIR, um sistema híbrido de Geração de Linguagem Natural que combina módulos determinísticos com LLMs baseados em transformadores (por exemplo, ChatGPT). O AGIR alcança alto recall (0,99), evita alucinações e reduz o tempo de escrita de relatórios em mais de 40%, automatizando a produção de relatórios abrangentes a partir de grafos de entidades em SOCs. Evoluindo de relatórios estáticos para suporte dinâmico à resposta a incidentes, Ismail et al. [8] propõem o Security Event Response Copilot (SERC), integrando RAG (*Retrieval-Augmented Generation*) para extração de dados de eventos e orientação de resposta baseada em LLM. O SERC evidencia aplicabilidade empresarial ao apoiar o ecossistema de SOC (*Security Operation Center*) de código aberto como alternativa econômica às soluções proprietárias.

Em sistemas de recuperação aprimorados por conhecimento, Kurniawan [10] discute limitações de LLMs em raciocínio complexo e consistência em múltiplas consultas, propondo o CyKG-RAG. O framework ancora os modelos em grafos de conhecimento (CVE, CAPEC (*Common Attack Pattern Enumeration and Classification*), MITRE ATT&CK (*Adversarial Tactics, Techniques, and Common Knowledge*)), mitigando alucinações ao fornecer relações semânticas estruturadas e unificando inteligência de ameaças estruturada e logs não estruturados via representações em grafo.

2.3 Automação de compliance e gestão de riscos

Salman, Creese e Goldsmith [19] criticam abordagens tradicionais de conformidade baseadas em documentação e defendem uma avaliação baseada em evidências, explorando o uso de LLMs para análise de dados não estruturados e monitoramento contínuo de conformidade, superando auditorias periódicas e ampliando cobertura organizacional.

Para TPRM (*Third-Party Risk Management*), Arora [1] integra GAI e RAG em um *framework* que processa políticas de terceiros em múltiplos formatos (por exemplo, PDFs) e incorpora supervisão humana por verificação cruzada via hiperlinks, equilibrando eficiência de automação e controle humano.

2.4 Desafios e implicações

Observou-se, na literatura, que persistem desafios técnicos relevantes quanto ao uso de GAI mas que existem mecanismos de contorno. Kurniawan [10] ressalta a propensão dos LLMs a introduzir informações irrelevantes/incorretas (alucinação) em conjuntos factuais; abordagens orientadas por conhecimento (CyKG-RAG) mitigam, mas não eliminam, questões de confiabilidade, viés, justiça e explicabilidade. Observam-se também lacunas de orientação prática para empresas de médio porte com restrições típicas de recursos, orçamento, expertise e carência de validação empírica ampla. Perrina et al. [16] reportam métricas quantitativas e Ismail et al. [8] demonstram integração prática, enquanto Salman, Creese e Goldsmith [19] e Arora [1] permanecem sobretudo conceituais.

Ainda assim, há potencial expressivo de ganhos: a redução de 40% no tempo de relatório no AGIR e a abordagem de custo viável do SERC sugerem ganhos de produtividade sem custos proporcionais de licenciamento, aspecto crítico para médias empresas. Por outro lado, a complexidade de manter grafos de conhecimento e infraestrutura LLM (*Large Language Model*)(CyKG-RAG) pode exceder a capacidade técnica de PMEs. Em pano de fundo, Karpatou [9] contextualiza a evolução de ameaças (*phishing* e engenharia social alimentados por IA, *malware* autônomo/auto-evolutivo, *deepfakes* e fraude de identidade), reforçando a urgência de mecanismos de defesa avançados, embora os caminhos práticos de adoção permaneçam desafiadores.

2.5 Evolução das abordagens anteriores

As aplicações atuais de GAI representam um avanço substancial sobre sistemas rígidos baseados em regras ou correspondência de palavras-chave, que não capturavam nuances regulatórias. SIEMs (*Security Information and Event Management*) tradicionais agregam e correlacionam eventos, mas carecem das capacidades de PLN (Processamento de Linguagem Natural) demonstradas por LLMs integrados a grafos de conhecimento.

Em TPRM, onde avaliações estáticas não acompanham requisitos regulatórios em evolução, abordagens ancoradas em RAG [1] oferecem atualização regulatória em tempo quase real, embora careçam, até o momento, de validação empírica abrangente.

2.6 Lacunas de pesquisa e direções futuras

Percebeu-se que uma lacuna importante na área e a ausência de validação empírica rigorosa dos frameworks GAI propostos em ambientes empresariais reais, incluindo benchmarks quantitativos de precisão, eficiência e custo-benefício frente a processos tradicionais. Estudos futuros devem priorizar experimentos controlados, avaliações sob regimes regulatórios dinâmicos e análises específicas para PMEs (restrições de recursos, escalabilidade, integração com legados).

Há necessidade de modelos de implantação simplificados e de pesquisas sobre segurança de IA (detecção/mitigação de viés, explicabilidade, governança), bem como sobre aceitação regulatória de evidências geradas por IA. Em síntese, a GAI tem potencial para aliviar substancialmente a dificuldade da geração manual de relatórios de conformidade sintetizando alertas de vulnerabilidades e logs para produzir relatórios abrangentes com menor propensão a erros humanos, mas sua adoção exige cuidados com dados, privacidade e confiança na acurácia das saídas.

2.7 Considerações

A revisão consolidou aplicações emergentes de GAI em inteligência de ameaças, recuperação ancorada em conhecimento e avaliação de conformidade baseada em evidências, com ganhos potenciais de eficiência especialmente relevantes para empresas de médio porte. Persistem, contudo, desafios de segurança e confiabilidade de IA, lacunas de implementação específicas para PMEs e insuficiência de validação empírica dos *frameworks*. Esses achados orientam a pesquisa aplicada deste trabalho, com foco na proposta de arquitetura, critérios de avaliação e mecanismos de mitigação de riscos do uso de GAI em compliance de cibersegurança.

3. Metodologia utilizada

A pesquisa adota um enfoque aplicado, qualitativo, exploratório e descritivo. No plano conceitual, parte de uma revisão sistemática da literatura, conforme descrito na seção anterior, que permitiu identificar o estado da arte em aplicações de IA Generativa em segurança e compliance, as principais abordagens utilizadas e lacunas que a proposta buscou endereçar.

No plano técnico, o trabalho segue um ciclo de projeto e avaliação que pode ser caracterizado em quatro etapas principais. Na primeira etapa, são definidos os requisitos funcionais e não funcionais da solução, a partir da análise do contexto de empresas de médio porte, dos desafios de escassez de profissionais e da natureza das demandas de compliance em segurança e proteção de dados. Na segunda etapa, é projetada uma arquitetura modular capaz de integrar fontes de dados de segurança, processá-las e utilizar modelos de IA Generativa para produzir relatórios orientados a personas. Na terceira etapa, essa arquitetura é implementada em um protótipo funcional, incluindo um fluxo preliminar em n8n para validar o encadeamento de atividades e uma implementação principal em Python para suportar maior flexibilidade e escalabilidade. Por fim, na quarta etapa, a solução é avaliada por meio de experimentos com dados simulados e reais e de uma avaliação qualitativa estruturada, utilizando um questionário com escala Likert.

Para a avaliação com usuários, este trabalho buscou representantes técnicos e um gestor da instituição utilizada para coleta de dados reais. Antes dessa etapa, já com a matriz de 18 relatórios gerados (3 personas x 6 modelos de IA) criou-se um agente de GAI orientado pelo contexto da metodologia de avaliação proposta. Com base no resultado da análise automatizada destacou-se o relatório gerado pelo modelo de GAI com melhor resultado para cada persona, estes foram os relatórios submetidos a avaliação dos especialistas. Cada participante recebeu relatórios gerados automaticamente por diferentes modelos de IA para sua persona específica e respondeu a um questionário baseado em escala Likert de cinco pontos (1 = discordo totalmente, 5 = concordo totalmente). As questões foram organizadas em três dimensões: qualidade linguística, alinhamento ao propósito e desempenho de IA. A partir das respostas, foram calculadas médias por dimensão, por persona e por modelo, de modo a possibilitar a construção de um ranking de desempenho e a análise comparativa dos modelos e da adequação da abordagem proposta.

4. Desenvolvimento da Solução

Esta seção descreve o desenvolvimento da solução proposta para automatizar a geração de relatórios de compliance em segurança da informação e proteção de dados. A implementação parte dos desafios identificados em empresas de médio porte escassez de profissionais especializados, necessidade de consolidar múltiplas fontes de evidências e exigência de relatórios adaptados a diferentes áreas corporativas e materializa uma arquitetura modular que integra SIEM, processamento de dados e agentes de IA Generativa.

A solução foi concebida inicialmente em um protótipo de validação arquitetural na plataforma de automação n8n[5] e, em seguida, reimplementada em Python, com um pipeline completo que vai da coleta de dados no Wazuh SIEM/XDR até a geração de relatórios personalizados para diferentes personas (CEO (*Chief Executive Officer*), TIC e Compliance Officer), incluindo uma matriz de testes para avaliação comparativa de múltiplos modelos de linguagem.

4.1 Arquitetura da solução

A arquitetura lógica da solução, ilustrada na

, foi projetada de forma modular, com três blocos principais: (i) Integração de Dados; (ii) Processamento de Dados; e (iii) Geração de Relatórios. A Integração de Dados é responsável por coletar informações de sistemas de gestão de vulnerabilidades Wazuh SIEM/XDR [6] e de a possiblidade de inclusão de informações frameworks de boas práticas e requisitos regulatórios, como normas de segurança e legislações de proteção de dados como contexto para a GAI. A camada de Processamento de Dados transforma esses dados brutos em estruturas consolidadas, extraindo vulnerabilidades, eventos de segurança, indicadores de postura de conformidade com dados de SCA (*Security Configuration Assessment*) e agregando estatísticas relevantes por ativo, política e severidade. A Geração de Relatórios, por sua vez, utiliza modelos de IA Generativa para produzir narrativas orientadas a diferentes perfis de audiência, mantendo um *template* base determinístico de relatório de compliance e variando conteúdo e tom de acordo com cada persona.

O uso da IA Generativa é sempre condicionado por um contexto cuidadosamente preparado, composto por: (i) um template base de relatório de compliance, que garante estabilidade estrutural; (ii) dados de vulnerabilidades, eventos e conformidade extraídos da API (*Application Programming Interface*) do Wazuh e de frameworks de segurança; e (iii) instruções específicas de linguagem e foco para cada persona. Esse contexto é construído de forma iterativa: os dados são pré-processados e formatados, em seguida injetados em prompts dinâmicos que instruem o modelo sobre a tarefa, e os resultados são avaliados e refinados a partir de múltiplos modelos, com observação visual e verificações

[5] Disponível em: https://n8n.io/
[6] Disponível em: https://wazuh.com/

automáticas de termos para mitigar problemas de fluência ou de conteúdo.

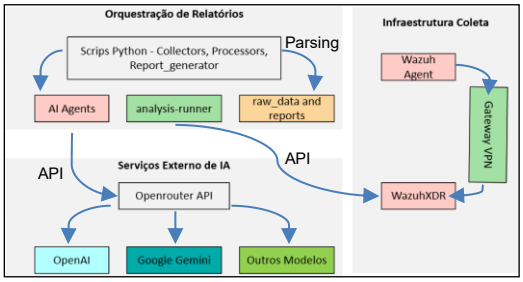

Figura 1. Arquitetura da Solução

4.2 Protótipo de testes em n8n

A primeira materialização prática da arquitetura foi construída na plataforma de automação n8n, escolhida por sua ampla adoção no mercado, modelo open source e grande variedade de conectores, o que permitiu acelerar a fase inicial de experimentação [12]. Neste protótipo implementou-se um pipeline completo desde a autenticação no Wazuh até a geração de relatórios, permitindo validar, com baixo custo, o conceito de automação de coleta, processamento e uso de agentes de IA Generativa.

Os resultados da etapa em n8n demonstraram a viabilidade técnica da abordagem: foi possível orquestrar a coleta paralela de dados, consolidar evidências de segurança e gerar relatórios executivos enriquecidos por análise de IA. Entretanto, tornaram-se evidentes limitações em relação a: (i) flexibilidade para manipular estruturas de dados complexas; (ii) desempenho ao processar volumes crescentes de informações; (iii) suporte a testes automatizados e tratamento avançado de exceções; e (iv) necessidade de bibliotecas mais especializadas do que aquelas disponíveis nativamente em nós JavaScript. Essas limitações motivaram a migração para uma implementação em Python, mantendo o aprendizado arquitetural obtido no protótipo.

4.3 Implementação em Python e ambiente de testes

A partir dos testes preliminares no n8n, a solução foi portada para um pipeline desenvolvido em Python, que oferece maior controle sobre estrutura de dados, testes, desempenho e integração com bibliotecas de IA. O ambiente de monitoramento é baseado no Wazuh SIEM/XDR instalado em nuvem privada, que atua como fonte centralizada de dados de segurança. Para os testes, utilizou-se um conjunto de seis servidores de homologação da Universidade do Vale do Itajaí, complementados por um servidor Linux propositalmente desatualizado e uma estação de trabalho do pesquisador. Esse arranjo permitiu criar um cenário com maior volume e diversidade de vulnerabilidades, utilizando a base NVD (*NIST Vulnerability Database*) e verificações de conformidade com boas práticas de configuração e reforço de segurança baseadas em OpenSCAP.

O Wazuh, segundo Stanković, Gajin e Petrović [21], é uma plataforma gratuita e de código aberto especializada em detecção de ameaças e monitoramento de segurança baseado em agentes. Ela coleta logs e métricas de *endpoints* heterogêneos e os envia a um servidor central para correlação e análise, oferecendo uma visão consolidada e em tempo (quase) real da infraestrutura. Comparativos com outras soluções de mercado mostram que o Wazuh oferece xcelente relação custo-benefício e interface amigável [21], fatores alinhados ao foco deste trabalho em empresas de médio porte.

A plataforma expõe uma API REST autenticada via token JWT (*JSON Web Token*), por meio da qual a solução extrai inventário de ativos, vulnerabilidades (CVEs) e resultados de SCA. Os parâmetros de conexão e acesso são definidos em um arquivo .env, incluindo variáveis como WAZUH_API_URL, WAZUH_API_USER, WAZUH_API_PASSWORD e flags de verificação de certificado (VERIFY_SSL). Variáveis adicionais permitem, quando necessário, o acesso direto ao Wazuh Indexer para consultas mais sofisticadas.

Já o acesso à IA Generativa é configurado via *OpenRouter*[7], também por variáveis em .env, como *OPENROUTER_API_KEY, OPENROUTER_API_BASE_URL e OPENROUTER_MODEL*, cuja escolha influencia custo, latência e qualidade do texto gerado. Parâmetros de geração de relatórios, como *GENERATE_TRANSITIONS_BY_AI* (para habilitar ou não parágrafos de transição), RENDER_PDF e *REPORT_OUTPUT_DIR* completam a configuração, permitindo ajustar o comportamento do pipeline sem alterar código.

Arquiteturalmente, a implementação em Python se organiza em cinco módulos principais: (i) coletores, responsáveis pela ingestão de dados brutos; (ii) processadores, que normalizam e enriquecem os dados; (iii) um módulo de consolidação e transição textual; (iv) os agentes de IA de geração de relatório; e (v) a camada de renderização, validação pós-render e persistência, a Figura 2 ilustra o fluxo de dados do pipeline de geração dos relatórios.

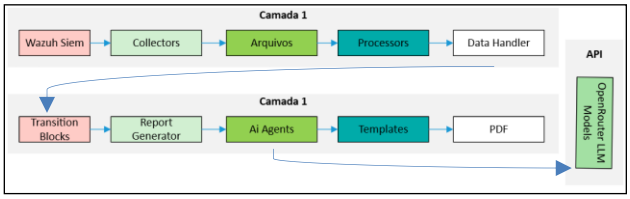

Figura 2. Fluxo de Dados do Pipeline

Os coletores gravam seus resultados em raw_data/. O módulo *inventory_collector* extrai o inventário de agentes do Wazuh, incluindo sistema operacional (distribuição, versão, arquitetura), endereços IP, datas de registro, estado de conexão e outros metadados, serializando o resultado em JSON estruturado. Já o módulo *data_collector* concentra-se em vulnerabilidades e SCA, realizando requisições ao *endpoint* de vulnerabilidades organizando os dados por agente de coleta e, por fim as consultas aos resultados de SCA. Esse módulo implementa *rate limiting*, já que o conjunto de dados pode ser bastante grande e a consulta exaurir os recursos da plataforma Wazuh, além de mecanismos de

[7] Disponível em: https://openrouter.ai/

checkpoint que permitem retomar a coleta em caso de interrupções, evitando reprocessamento desnecessário.

Na sequência, o módulo processador, transforma os dados brutos em estruturas otimizadas para consumo pelos agentes de relatório. O *inventory_processor* enriquece o inventário, aplica mapeamentos opcionais de nomes de host, classifica e agrupa agentes e gera arquivos processados (*processed_inventory.json*). O data_processor também trata dos dados de vulnerabilidades e SCA: carrega os dados coletados, enriquece CVEs (*Common Vulnerabilities and Exposures*) com informações de CVSS (score, vetor, severidade), agrupa e prioriza vulnerabilidades (por exemplo, top 10 críticas), processa resultados de SCA por política de conformidade, calcula taxas de conformidade, identifica verificações mais problemáticas e produz estatísticas agregadas persistidas em arquivos JSON/CSV.

Para que o texto produzido por cada agente de IA pudesse apresentar um resultado mais fluido, um módulo intermediário, como build_transition.py, foi desenvolvido com objetivo de empacotar blocos de transição textual (por exemplo, entre seções de sumário executivo, análise de vulnerabilidades e evidências) para serem usados opcionalmente por um agente de transições, melhorando a fluidez narrativa do relatório.

4.4 Geração de relatórios e matriz de testes

A partir dos dados processados, o pipeline é capaz de gerar relatórios adaptados a diferentes personas. O comando de orquestração (run_full_pipeline ou report_generator) recebe parâmetros que indicam a persona-alvo (por exemplo, ceo, tic ou compliance_officer), se a coleta e o processamento devem ser reexecutados ou reutilizados de execuções anteriores, e se a versão final em PDF deve ser gerada.

Para o perfil CEO, os templates de prompt enfatizam impacto de negócio, riscos estratégicos, indicadores agregados e recomendações priorizadas, resultando em relatórios com sumário executivo, KPIs de segurança, análise de impacto de vulnerabilidades críticas e um roadmap de remediação com foco em custo-benefício. Para o perfil TIC (CISO (*Chief Information Security Officer*)/Diretor de TI), o relatório aprofunda aspectos técnicos e de governança, detalhando vulnerabilidades por categoria, avaliação da postura de conformidade, mapeamento de riscos por ativo crítico e recomendações específicas de hardening e patching. Já para o Compliance Officer, o conteúdo enfatiza aderência a políticas e frameworks, lacunas de conformidade, evidências de controles e recomendações orientadas a requisitos regulatórios.

Independente da persona, a geração segue a mesma arquitetura multiagente: um conjunto de agentes de IA especializados recebe o consolidated_data e gera seções específicas do relatório. Entre eles, destacam-se: ReportIntroductionAgent (contextualização e objetivo do relatório), EnvironmentCharacterizationAgent (caracterização do ambiente monitorado), VulnerabilityAnalysisAgent (análise de CVEs e riscos associados), SCAAnalysisAgent (postura de conformidade), ExecutiveSummaryAgent (síntese de descobertas e riscos prioritários) e ConclusionAgent (conclusões e recomendações finais). Opcionalmente, um TransitionGenerator produz parágrafos de ligação entre seções, reforçando coesão e fluidez textual.

Após a geração textual, o conteúdo é injetado em templates HTML baseados em Jinja2[8], formando o relatório em HTML. Um PostRenderValidator verifica a existência de placeholders não resolvidos, inconsistências ou trechos suspeitos (por exemplo, potenciais alucinações), podendo registrar e descartar partes problemáticas num arquivo de validação (post_render_validation.json). Se habilitado, o HTML é convertido em PDF. Cada execução gera uma pasta em generated_reports/ contendo o HTML final, o PDF (quando solicitado), arquivos de contexto usados pelos agentes (data_context_for_xxx.json), registros de validação pós-render e uma pasta evidence/ com os artefatos brutos de coleta e processamento, garantindo rastreabilidade e auditabilidade.

Por fim, a solução implementa uma matriz de testes para avaliar comparativamente diferentes modelos de linguagem em múltiplas personas. Um script específico (generate_matrix_reports.py) automatiza a geração em lote de relatórios para todas as combinações de três personas padrão (CEO, Compliance Officer e TIC) com seis modelos de linguagem. Entre eles, há modelos gratuitos, alternativas baseadas em grandes modelos de código aberto e modelos "premium" de maior capacidade. O resultado é uma matriz de 18 relatórios gerados a partir dos mesmos dados consolidados de segurança.

A execução completa da matriz, produziu esses 18 relatórios em aproximadamente 49 minutos, o que corresponde a cerca de 2,7 minutos por relatório. Observou-se que a maior parte desse tempo está associada ao tempo de resposta dos modelos de IA, que varia conforme o tipo (modelos simples vs. de raciocínio profundo) e o perfil de custo (modelos "premium" tendem a ser mais rápidos). Para garantir reprodutibilidade e isolar o comportamento dos modelos, a coleta e o processamento de dados foram executados apenas uma vez no início do processo em lote, com tempo médio de cerca de 50 segundos no ambiente de testes.

A cabo desta etapa de desenvolvimento observou-se que a arquitetura modular implementada em Python é capaz de operacionalizar, de ponta a ponta, a solução proposta: coletar dados de segurança em um SIEM de mercado, processá-los em uma estrutura consolidada, gerar relatórios adaptados a diferentes personas com suporte de IA Generativa e fornecer um arcabouço experimental (matriz de testes) para avaliar comparativamente múltiplos modelos de linguagem em um cenário realista de compliance em segurança da informação.

5. Resultados e Discussão

A avaliação da solução proposta envolve tanto aspectos qualitativos, relacionados à percepção dos usuários sobre os relatórios gerados, quanto aspectos quantitativos, relacionados à eficiência, escalabilidade e comparação entre modelos de IA. No componente qualitativo, a aplicação do questionário com escala

[8] Disponível em: https://jinja.palletsprojects.com/en/stable/

Likert foi estruturado em três dimensões: qualidade linguística, alinhamento ao propósito e desempenho da IA. Tal abordagem permitiu capturar impressões de participantes representantes das personas de CEO, Compliance Officer/jurídico e TIC. Na dimensão de qualidade linguística, as questões abordaram clareza, coerência, estrutura do texto, adequação do vocabulário ao público-alvo e tom utilizado. Na dimensão de alinhamento ao propósito, foram avaliados a relevância do conteúdo para as responsabilidades de cada persona, a utilidade e possibilidade de implementação das recomendações, a forma de comunicação de riscos e a adequação de elementos como sumário executivo e artefatos de apoio (tabelas, gráficos e listas). Já na dimensão de desempenho da IA, as perguntas investigaram a consistência factual entre relatórios, a fidelidade aos dados de entrada, a profundidade da análise produzida, o equilíbrio entre originalidade e repetição de padrões textuais e a presença de erros factuais, gramaticais ou de formatação. Em termos quantitativos, apresenta-se, na

Tabela 2, um ranking geral de desempenho por modelo, construído a partir das médias das três dimensões avaliadas em todas as personas.

Tabela 2. Ranking Geral de Desempenho por Modelo

Posição	Modelo	CEO	Compliance Officer	TIC	Média Geral
1º	Google Gemini 2.5 Flash Lite	4.94	5.00	5.00	4.98
2º	X.AI Grok 4.1 Fast	4.94	4.94	4.94	4.94
3º	DeepSeek Chat v3	4.43	5.00	4.92	4.78
4º	Google Gemini 3 Pro Preview	4.14	4.44	4.29	4.29
5º	Z.AI GLM 4.5 Air Free	4.14	4.14	4.22	4.17
6º	OpenAI GPT-OSS 20B Free	3.43	3.51	3.43	3.46

Esse ranking mostra que há um grupo de modelos que se destaca com médias globais mais altas e desempenho consistente em qualidade linguística, alinhamento ao propósito e desempenho de IA. Outros modelos, embora funcionais, apresentaram desempenho inferior em pelo menos uma dimensão, como textos menos claros ou menor adaptação ao perfil do leitor. A Tabela 3 detalha as médias por dimensão, permitindo observar, por exemplo, que determinados modelos se sobressaem na clareza da escrita, enquanto outros obtêm melhor avaliação em alinhamento semântico ao propósito do relatório.

Tabela 3. Performance Média por Dimensão de Avaliação

Modelo	A (Linguística)	B (Propósito)	C (IA)	Consistência Cross-Persona
Gemini 2.5 Flash Lite	5.00	5.00	4.93	★★★★★ Excelente
Grok 4.1 Fast	5.00	5.00	4.80	★★★★★ Excelente
DeepSeek Chat v3	4.83	4.73	4.80	★★★★ Boa
Gemini 3 Pro Preview	4.42	4.20	4.27	★★★ Regular
GLM 4.5 Air Free	4.25	4.07	4.20	★★★★ Boa
GPT-OSS 20B Free	3.75	3.27	3.40	★★ Fraca

Com base nas repostas, a análise por persona, revela variações interessantes. Para a persona CEO, os modelos mais bem avaliados produziram relatórios concisos, com ênfase em riscos estratégicos, continuidade de negócios e recomendações de alto nível, evitando jargão técnico excessivo. Modelos com desempenho inferior, por outro lado, tendiam a incluir detalhes técnicos supérfluos para esse público, diluindo a mensagem principal. Para a persona Compliance Officer, os melhores modelos geraram relatórios que estruturavam informações em torno de obrigações e controles, destacando potenciais lacunas de conformidade e sugerindo ações de mitigação compatíveis com o arcabouço regulatório considerado. Modelos de menor desempenho nessa persona produziam textos mais genéricos, com menor conexão explícita entre eventos específicos e requisitos normativos. Para a persona TIC, os modelos de melhor avaliação conseguiram oferecer descrições mais detalhadas de incidentes e vulnerabilidades relevantes, além de indicar prioridades e diretrizes iniciais de remediação, enquanto modelos menos ajustados tenderam a permanecer em um nível de abstração alto demais, exigindo que o analista recorresse novamente ao SIEM para obter o contexto técnico necessário.

Apesar de se considerar a avaliação comparativa entre modelos ponto importante, este trabalho propôs e desenvolveu uma arquitetura de geração automatizada de relatórios de segurança, a qual é fator de grande importância no objetivo geral definido o que demonstrou eficiência notável. A análise dos resultados, demonstrou a eficiência da arquitetura de geração. Em termos de desempenho técnico, a solução foi capaz de gerar relatórios em menos de 3 minutos a partir dos dados consolidados, o que contrasta com o tempo significativamente maior necessário para a redação manual. A arquitetura mostrou-se eficiente em uso de recursos computacionais, uma vez que a maior parte do esforço pesado é externalizado para o provedor de IA, e o processamento local fica para a tarefa de normalização e agregação de dados. A capacidade de reutilizar o mesmo conjunto de dados consolidados para gerar relatórios para múltiplas personas, sem retrabalho significativo, confirma a hipótese de que a abordagem aumenta a escalabilidade da atuação dos profissionais de segurança e compliance.

Os resultados também evidenciam a importância de uma engenharia cuidadosa de prompts. Mesmo entre os modelos de melhor desempenho, os relatórios só foram avaliados como bem alinhados ao propósito quando os templates especificavam com clareza o papel do modelo, o público-alvo e a estrutura desejada. Esse achado reforça conclusões da literatura sobre o uso de LLMs em tarefas específicas: a qualidade da saída depende não apenas das capacidades intrínsecas do modelo, mas também da forma como o problema é formulado[4].

Por outro lado, é importante destacar os riscos e limitações. Em cenários em que o contexto fornecido é incompleto ou ambíguo, há risco de que a IA produza extrapolações não suportadas pelos dados, fenômeno conhecido como *hallucination*. Essa possibilidade é particularmente sensível em relatórios de compliance, em que afirmações imprecisas podem ter implicações regulatórias. Por isso, a solução é explicitamente projetada como apoio à elaboração de relatórios, e não como substituto integral da revisão humana. Em linhas com as preocupações de Matas e Keegan [11] e Medeiros e Codignoto [13] sobre integridade e cultura de conformidade, este trabalho recomenda que relatórios gerados para fins regulatórios passem por revisão de profissionais de segurança e compliance, inserindo a IA como elemento de automação dentro de um processo governado.

Em síntese, a análise dos resultados indica que a solução proposta é capaz de reduzir substancialmente o esforço de produção de relatórios, aumentar a consistência documental, adaptar a comunicação a diferentes personas e manter um nível aceitável de confiabilidade, desde que acompanhada por supervisão humana e boas práticas de governança de dados e modelos.

6. Conclusão e Trabalhos Futuros

Este artigo, baseado no resultado do trabalho de uma dissertação de mestrado, apresentou uma investigação aplicada sobre o uso de Inteligência Artificial Generativa para automatizar a geração de relatórios de compliance em segurança da informação e proteção de dados em empresas de médio porte. A partir de uma revisão sistemática da literatura, do desenho de uma arquitetura modular integrando um SIEM a modelos de IA Generativa, da implementação de um protótipo funcional em Python e de testes preliminares em n8n, foram obtidas evidências de que a abordagem proposta contribui para otimizar a eficiência, a escalabilidade e a confiabilidade de processos de documentação de compliance.

A avaliação com usuários, estruturada em torno de um questionário com escala Likert em três dimensões (qualidade linguística, alinhamento ao propósito e desempenho de IA), permitiu comparar diferentes modelos de IA e verificar que há um conjunto de modelos que produz relatórios com boa clareza, forte aderência às necessidades de personas específicas e consistência razoável com os dados de base. A análise por persona revelou que relatórios gerados para executivos, profissionais de compliance e equipes técnicas podem, a partir do mesmo conjunto de dados consolidados, enfatizar aspectos distintos – riscos estratégicos, obrigações regulatórias ou detalhes técnicos – desde que os prompts sejam cuidadosamente desenhados. A análise da eficiência da arquitetura indicou que a solução é capaz de gerar relatórios em tempo reduzido, com uso moderado de recursos computacionais e boa capacidade de reutilização de dados para múltiplas saídas.

Os achados sustentam a hipótese geral de que uma solução baseada em IA Generativa pode apoiar empresas de médio porte na manutenção de programas de segurança da informação e proteção de dados mais robustos e bem documentados, apesar da escassez de profissionais. Ao facilitar a produção sistemática de evidências e relatórios de conformidade, a proposta contribui para reduzir riscos operacionais e regulatórios e para fortalecer a resiliência das infraestruturas digitais que suportam serviços e negócios em ambientes urbanos, em consonância com o Objetivo de Desenvolvimento Sustentável 11.

Como trabalhos futuros, o estudo sugere a integração da solução com plataformas de GRC (Governança, Risco e Compliance), de modo a ampliar a automação para além da geração de relatórios, alcançando registro de riscos, acompanhamento de planos de ação e monitoramento de controles. Aponta também para a necessidade de avaliações em larga escala, em ambientes de produção heterogêneos, que permitam medir o impacto da solução em diferentes setores e regimes regulatórios. Por fim, recomenda o aprofundamento em técnicas de engenharia de prompts e gestão de contexto, bem como o desenvolvimento de mecanismos de explicabilidade e rastreabilidade que facilitem auditorias e aumentem a confiança de reguladores e partes interessadas na adoção de IA Generativa em processos de compliance.

REFERÊNCIAS

[1] Arora, S. 2024. Revolutionizing Third Party Risk Management Using Generative AI and RAG. *2024 First International Conference on Data, Computation and Communication (ICDCC)* (2024).

[2] Disterer, G. 2013. ISO/IEC 27000, 27001 and 27002 for Information Security Management. *Scientific Research Publishing*. 4, 2 (Jan. 2013), 92–100. https://doi.org/10.4236/jis.2013.42011.

[3] European Council 2023. Cybersecurity: how the EU tackles cyber threats.

[4] George, A.S. et al. 2023. ChatGPT and the Future of Work: A Comprehensive Analysis of AI's Impact on Jobs and Employment. *Partners Universal International Innovation Journal*. 1, 3 (2023), 154–186. https://doi.org/10.5281/zenodo.8076921.

[5] Harvard Business Review 2019. The Public-Private Partnership That's Working to Make New York City a Global Hub of Cybersecurity Talent.

[6] IBM 2022. IBM Security Incident Responder Study.

[7] (ISC)^2 2022. Cybersecurity Workforce Study.

[8] Ismail et al. 2025. Enhancing Security Operations Center: Wazuh Security Event Response with Retrieval-Augmented-Generation-Driven Copilot. *Sensors*. 25, 3 (2025), 870.

[9] Karpatou, P.A. 2025. The evolution of cybersecurity threats & the rise of artificial intelligence. *Bachelor dissertation*. (2025).

[10] Kurniawan, K. 2024. CyKG-RAG: Towards knowledge-graph enhanced retrieval augmented generation for cybersecurity. *Ceur Workshop Proceedings* (2024).

[11] Matas, S.D. and Keegan, B.J. 2020. Challenges in Addressing Information Security Compliance in Healthcare Research: The Human Factor. *Science Publishing Group*. 5, 2 (Jan. 2020), 25–25. https://doi.org/10.11648/j.ajomis.20200502.12.

[12] McFeetors, J. and Pant, T. 2022. *Rapid Product Development with n8n*. Packt Publishing.

[13] Medeiros, M.L. and Codignoto, R. 2022. Governança, integridade e resultados caminham juntos. 3, 1 (Oct. 2022). https://doi.org/10.37497/regov.v3i1.30.

[14] Oreyomi, M. and Jahankhani, H. 2022. Challenges and Opportunities of Autonomous Cyber Defence (ACyD) Against Cyber Attacks. *Blockchain and Other Emerging Technologies for Digital Business Strategies*. Springer. 239–269.

[15] Pascoe, C.E. 2023. Public Draft: The NIST Cybersecurity Framework 2.0. *National Institute of Standards and Technology*. (2023).

[16] Perrina, F. et al. 2023. AGIR: Automating Cyber Threat Intelligence Reporting with Natural Language Generation. *2023 IEEE International Conference on Big Data (BigData)* (2023), 3053–3062.

[17] PRISMA 2020. PRISMA Flow Diagram.

[18] Sacomano Neto, M. and Escrivão Filho, E. 2000. Estrutura organizacional e equipes de trabalho: estudo da mudança organizacional em quatro grandes empresas industriais. *Gestão & produção*. 7, (2000), 136–145.

[19] Salman, A. 2024. Position Paper: Leveraging Large Language Models for Cybersecurity Compliance. *Proceedings 9th IEEE European Symposium on Security and Privacy Workshops Euro S and Pw 2024* (2024).

[20] Singh, T. et al. 2023. Stress in the cybersecurity profession: a systematic review of related literature and opportunities for future research. *Organizational Cybersecurity Journal: Practice, Process and People*. (2023). https://doi.org/10.1108/ocj-06-2022-0012.

[21] Stanković, S. et al. 2022. A review of Wazuh tool capabilities for detecting attacks based on log analysis. *No Nama Agent Integrity File Added Delete Modified*. 1, (2022).

[22] TechCrunch+ 2023. Google brings generative AI to cybersecurity.

[23] Vaswani, A. et al. 2017. Attention is all you need. *Advances in neural information processing systems*. 30, (2017). https://doi.org/10.48550/arXiv.1706.03762.