GT-LFI: Anonimização, Classificação e Aprendizagem Gamificada a partir de Incidentes Reais

Rodrigo S. Miani[2], Diego Kreutz[1], Silvio E. Quincozes[1,2], Leandro M. Bertholdo[3], Rafael D. Araújo[2], Felipe N. Dresch[1], Felipe H. Scherer[1], João P. R. Esteves[2], Sebastião A. de Jesus Filho[2], Alvaro S. Santos[2]

[1]Universidade Federal do Pampa (UNIPAMPA) – Brasil

[2]Faculdade de Computação – Universidade Federal de Uberlândia (UFU) – Brasil

[3]Universidade Federal do Rio Grande do Sul (UFRGS) – Brasil

miani@ufu.br, diegokreutz@unipampa.edu.br, silvioquincozes@unipampa.edu.br, berthold@penta.ufrgs.br, rafael.araujo@ufu.br, felipedresch.aluno@unipampa.edu.br, felipescherer.aluno@unipampa.edu.br, joao.esteves@ufu.br, sebastiao@ufu.br, alvaro.santana@ufu.br

Modalidade: Código Aberto

Abstract. *Incident-response training is often separated from operational practice, while real tickets cannot be reused directly because they contain sensitive data. This paper presents GT-LFI, an open-source ecosystem that turns real incidents into reusable knowledge through an end-to-end flow: structured anonymization, AI-assisted categorization and playbook generation, human validation, and conversion of validated cases into gamified learning paths. The ecosystem combines four interoperable tools and was evaluated with residents of the Hackers do Bem program. Results show practical use of lessons, quizzes and learning paths, positive perceived usefulness, and descriptive learning gains, while also revealing usability and instrumentation limitations. The demonstration will reproduce the complete journey from a raw incident to a validated educational path.*

Resumo. *A formação em resposta a incidentes ainda é frequentemente dissociada da prática operacional, enquanto tíquetes reais não podem ser reutilizados diretamente por conterem dados sensíveis. Este artigo apresenta o GT-LFI, ecossistema de código aberto que transforma incidentes reais em conhecimento reutilizável por um fluxo fim-a-fim: anonimização estruturada, categorização e geração de playbooks assistidas por IA, validação humana e conversão de casos validados em trilhas gamificadas. A solução integra quatro ferramentas interoperáveis e foi avaliada com residentes do programa Hackers do Bem. Os resultados mostram uso efetivo de lições, quizzes e trilhas, utilidade percebida e ganhos descritivos de aprendizagem, além de limitações de usabilidade e instrumentação. A demonstração reproduzirá a jornada completa, do incidente bruto à trilha educacional validada.*

1. Introdução

Responder a incidentes de segurança exige interpretar evidências incompletas, priorizar riscos, documentar decisões e executar ações de contenção e recuperação sob pressão. *playbooks* reduzem a variabilidade desse processo ao explicitar procedimentos e critérios de escalonamento, mas sua elaboração e manutenção demandam tempo de especialistas [Nelson et al. 2025]. Ao mesmo tempo, a formação tradicional tende a privilegiar conteúdos estáticos e exercícios descontextualizados, distantes dos tíquetes, logs e decisões encontrados em PoPs, CSIRTs e SOCs [Alnajim et al. 2023].

Incidentes históricos constituem uma fonte valiosa de casos práticos, porém frequentemente contêm endereços IP, nomes, domínios, URLs e detalhes de infraestrutura. Seu uso educacional depende, portanto, de mecanismos que protejam a privacidade sem destruir as relações semânticas necessárias à análise. Depois de preparados, esses registros ainda precisam ser categorizados, associados a respostas consistentes, revisados por pessoas e convertidos em atividades adequadas a diferentes competências.

Esse equilíbrio entre privacidade e utilidade permanece um problema aberto: pseudonização não elimina por si só o risco de reidentificação, enquanto remoções agressivas podem comprometer tarefas posteriores [Gadotti et al. 2024]. Abordagens de NLP mostram que mecanismos baseados em reconhecimento de entidades preservam melhor a utilidade analítica, ao passo que substituições generativas produzem textos mais naturais [Yermilov et al. 2023]. No eixo educacional, estudos recentes associam gamificação a maior engajamento, mas ressaltam que o efeito depende do perfil dos aprendizes, da coerência das recompensas e do desenho das atividades [dos Santos et al. 2025, Meng et al. 2024].

O GT-LFI (*Learn From Incidents*) foi desenvolvido no Programa P&D Hackers do Bem para conectar essas etapas em um único ciclo operacional e pedagógico. A contribuição do trabalho não é apenas uma interface isolada, mas um ecossistema composto por: (i) ANON-LFI, para preparação e pseudonimização de incidentes; (ii) AutoClass-LFI, para categorização e geração assistida de *playbooks*; (iii) Sistema PoP, para triagem, correção e validação humana; e (iv) Hackers do Bem, para autoria e consumo de trilhas gamificadas. O conhecimento validado retorna à base e pode apoiar novos casos, formando um ciclo humano-no-loop.

Este artigo descreve a arquitetura e as funcionalidades das ferramentas, relata as avaliações conduzidas em 2026 e detalha a demonstração planejada para o Salão de Ferramentas. A modalidade é Código Aberto e o ponto de entrada dos artefatos é https://github.com/rsmiani/gt-lfi-rnp.

2. Ecossistema GT-LFI

2.1. Fluxo operacional e pedagógico

A Figura 1 apresenta o fluxo do ecossistema. Uma fonte parceira fornece incidentes em formatos estruturados ou semiestruturados, como planilhas, mensagens e exportações de tíquetes. O ANON-LFI detecta atributos sensíveis, aplica pseudonimização consistente e produz um registro normalizado. O AutoClass-LFI usa esse registro para sugerir categoria, subcategoria, resumo técnico e ações de resposta; quando aplicável, produz um *playbook* em *Markdown* ou *Ansible*.

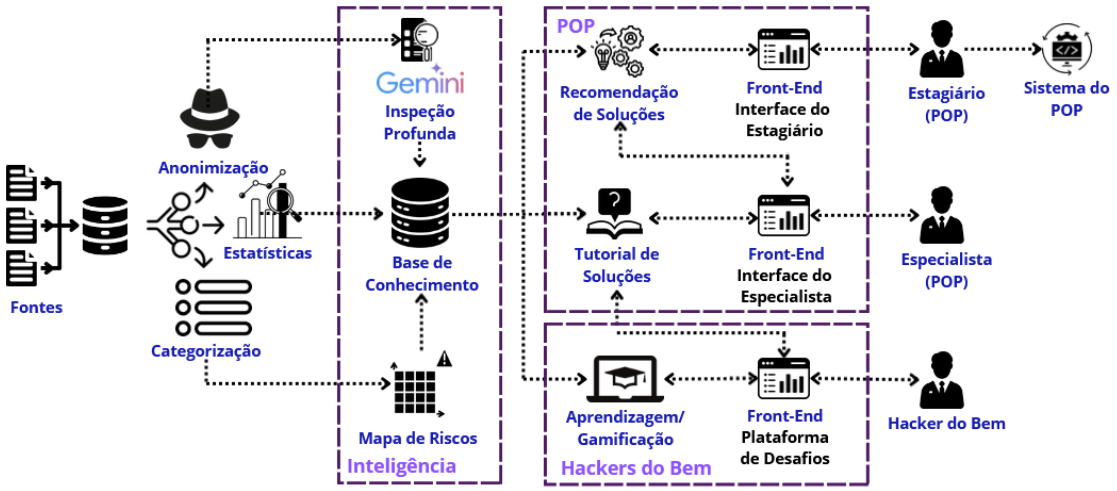

Figura 1. Arquitetura integrada e papéis humanos no ecossistema GT-LFI.

As sugestões não são tratadas como decisão final. No Sistema PoP, um estagiário realiza a primeira análise e pode confirmar ou corrigir a categorização e as ações. Um analista experiente revisa o resultado, registra sua decisão e aprova o caso. Esse processo preserva governança humana sobre saídas probabilísticas e cria exemplos curados para o ciclo de aprendizagem. Por fim, casos aprovados podem ser selecionados no Hub de Trilhas e transformados em lições, *quizzes*, desafios e exercícios da plataforma Hackers do Bem.

O mesmo fluxo atende a dois objetivos complementares. No eixo operacional, apoia padronização, triagem e transferência de conhecimento. No eixo educacional, aproxima aprendizes de situações autênticas sem exposição indevida de dados. A integração é mais relevante que cada módulo isolado: o incidente percorre uma cadeia rastreável até se tornar conteúdo, e o feedback humano pode corrigir o conhecimento usado em execuções posteriores.

2.2. Componentes e responsabilidades

Os quatro módulos são implantáveis separadamente, mas compartilham contratos e identificadores: o ANON-LFI produz o incidente preparado; o AutoClass-LFI acrescenta classificação, justificativa e *playbook* candidato; o PoP registra correções e aprovação; e o Hackers do Bem converte o caso curado em trilhas, atividades e indicadores de progresso.

3. Arquitetura e Implementação

3.1. Contratos, isolamento e rastreabilidade

A integração combina chamadas HTTP autenticadas e documentos persistidos no Firebase. O ANON-LFI envia lotes ao endpoint `POST /api/incidents/upload`, usando `x-api-key`; o AutoClass-LFI responde imediatamente e executa categorização e geração de *playbooks* em segundo plano. O documento resultante transita pelos estados `received`, `categorizing`, `generating_playbooks`, `completed` ou `error`.

O mesmo `ticket_id` é usado no texto anonimizado, no índice vetorial e nas filas do PoP.

A Tabela 1, no Apêndice A, resume tecnologias e contratos verificados nos repositórios. Cada componente possui imagem Docker e variáveis próprias; essa separação permite executar o fluxo completo ou usar as ferramentas de forma independente.

Cada artefato produzido pelo AutoClass-LFI recebe metadados de proveniência com versão do esquema, ambiente, origem, identificadores, coleções *Firebase*, índice *Pinecone* e modelo de IA. Ambientes de produção e de avaliação usam coleções e índices distintos. O PoP grava feedback de estagiário e especialista sem substituir automaticamente a classificação original; ao consultar um caso, a versão concluída prevalece sobre a versão ainda disponível, preservando a trilha de auditoria.

3.2. ANON-LFI: preparação segura dos incidentes

O ANON-LFI recebe arquivos de incidentes e permite selecionar as colunas que devem ser processadas, preservadas ou removidas. No repositório atual, uma API Django/DRF trata lotes CSV de até 4,5 MB, converte cada linha em tíquete e aplica o Microsoft Presidio em português. O processamento ocorre em lotes de 32 textos, com limiar de confiança 0,6; reconhecedores inadequados ao domínio e a entidade `DATE_TIME` são excluídos para reduzir transformações indevidas. Sua estratégia combina regras para entidades estruturadas e reconhecimento contextual para trechos livres. Entre os elementos tratados estão IPv4/IPv6, e-mails, domínios, URLs, nomes de usuário e referências institucionais. A pseudonimização determinística preserva correlações úteis sem revelar o valor original [Bandel et al. 2025].

Cada bloco recebe um `ticketID` existente ou um identificador `INC-AAAAMMDD-*`, e os blocos são separados por `###`. O cliente do AutoClass-LFI adota *timeout* de 60 segundos e até três tentativas com espera exponencial para falhas transitórias. Além do conteúdo transformado, o módulo registra volume e entidades detectadas. Dados de reidentificação autorizada permanecem separados e sujeitos a controle de acesso. Essa separação permite que os módulos seguintes operem apenas sobre registros preparados.

3.3. AutoClass-LFI: categorização e geração de *playbooks*

O AutoClass-LFI recebe o texto anonimizado e produz saídas estruturadas. A implementação em Next.js e Express separa provedor de modelos, recuperação de contexto, esquemas de validação e apresentação. Na versão avaliada, o motor utilizou Gemini 2.0 Flash; a versão atual integra *Gemini*, *embeddings* de 768 dimensões, *Pinecone* e *Firebase*, mantendo índices distintos para uso interativo e para o fluxo automatizado do ANON-LFI. Técnicas de engenharia de prompts foram avaliadas anteriormente para categorização de incidentes [Severo et al. 2025], e um cenário *few-shot* foi investigado em trabalho complementar [Jesus Filho et al. 2025].

O sistema produz classificações CERT, NIST e livre, além de causa provável, síntese e duas respostas: uma versão legível por pessoas e outra em *Ansible*. Para reduzir desvios, os prompts combinam papel de analista, definições explícitas das taxonomias, regras que exigem cada ID exatamente uma vez e justificativa por decisão. Categorização usa temperatura 0,1; recomendações usam 0,7; as respostas têm limite de 4.000 tokens. Um

esquema Zod rejeita campos malformados, e a pós-validação verifica existência, unicidade e ordem dos IDs. Como LLMs podem produzir conteúdo plausível sem fundamentação [Zhang et al. 2025], toda recomendação permanece provisória até a validação humana.

Na avaliação técnica disponível, a classificação apresentou acurácia global de 44,3% e macro-F1 de 0,474 em um conjunto desbalanceado de 185 amostras. Algumas classes atingiram F1 de 0,840 e 0,857, enquanto classes raras tiveram desempenho baixo. Esses resultados justificam a arquitetura humano-no-loop e orientam o uso de mais exemplos, balanceamento e calibração, em vez de automação autônoma da resposta.

3.4. Sistema PoP e plataforma Hackers do Bem

O Sistema PoP oferece visões distintas para estagiários e especialistas. A primeira apresenta incidente, classificações CERT/NIST/LLM e *playbook* candidato; o estagiário pode assumir o tíquete, editar campos, justificar mudanças e salvar progresso. O especialista acessa a versão original e a revisada, refina a decisão e encerra o caso. As filas geral, estagiário, especialista, finalizados e "meus tíquetes" são materializadas em coleções *Firestore*; autenticação e papel são verificados separadamente, e a posse do tíquete é registrada pelo UID.

A plataforma Hackers do Bem transforma casos validados em experiências de aprendizagem. No Hub de Trilhas, instrutores selecionam um incidente curado e usam *Gemini* para esboçar estruturas, textos, quizzes ou exercícios de programação, com saída revisável antes da publicação. Trilhas contêm módulos e lições, com estimativas distintas para teoria, avaliação e prática. Estudantes acompanham progresso, XP, níveis, medalhas e ranking; os pontos vêm da conclusão de lições, quizzes e exercícios, com valor definido pelo instrutor, e determinam o nível N pela fórmula $\lfloor 100(N-1)^{1,5} \rfloor$, com $N \geq 1$ ($N = 1$ exige zero XP). O expoente 1,5 torna cada nível mais custoso (e.g., $N = 10$ exige 2.700 pontos), de modo que a progressão é formalmente aberta, mas com teto prático imposto pela curva. Cada nível também concede moedas ($30N + 20$, mais uma a cada dez XP). *Firebase* cuida de autenticação e persistência, e o PostHog registra eventos pedagógicos e sinais de fricção.

4. Avaliação e Resultados

4.1. Desenho multi-método

O MVPv1 foi avaliado com residentes do Hackers do Bem em duas rodadas, realizadas de 5 a 14 e de 19 a 23 de janeiro de 2026. A coleta combinou três fontes: (i) questionários autoadministrados em escala Likert de cinco pontos, precedidos por TCLE; (ii) telemetria do PostHog; e (iii) avaliações livres enviadas por e-mail. Foram consolidadas 35 respostas válidas e nove e-mails, sendo dois no corpo da mensagem e sete com anexos. O PoP, o AutoClass-LFI e a plataforma educacional foram disponibilizados aos participantes. O ANON-LFI, devido ao seu papel administrativo e ao tratamento de dados sensíveis, foi avaliado separadamente.

O questionário PoP/AutoClass adotou construtos do TAM, como utilidade, facilidade de uso e intenção de uso [Davis 1989], enquanto o instrumento do Hackers do Bem foi inspirado na UTAUT, contemplando expectativa de desempenho e esforço, influência social, condições facilitadoras, risco e intenção comportamental [Venkatesh et al. 2003]. Cada sistema possuía um projeto independente no PostHog. Foram coletados eventos

nativos, como visualização, saída de página, *dead click* e *rage click*, além de eventos de negócio, como assumir e finalizar tíquetes, concluir lições, submeter quizzes e adquirir itens [PostHog 2026]. Essa triangulação permite combinar percepção declarada, comportamento observado e relatos contextuais.

Após a priorização de 31 melhorias, uma nova versão educacional foi avaliada em junho. Nessa etapa, 14 pessoas realizaram o pré-teste, seis o pós-teste e cinco concluíram ambos; nove responderam ao instrumento de aceitação. Como os grupos diferem em tamanho e composição, os resultados permanecem descritivos e devem ser vistos como evidências preliminares.

4.2. Resultados técnicos por módulo

As evidências dos módulos são complementares. No ANON-LFI, 100 tíquetes sintéticos continham 516 entidades anotadas. A ferramenta detectou 358, das quais 350 eram verdadeiros positivos e oito falsos positivos; os 166 falsos negativos concentraram-se sobretudo em organizações, cuja forma lexical é variável. A precisão de 0,9777 indica baixo risco de transformar trechos válidos, mas a revocação de 0,6783 exige ampliar reconhecedores e cobertura. O F1 foi 0,8009.

No AutoClass-LFI, 185 tíquetes reais previamente validados foram classificados segundo a taxonomia NIST. A acurácia foi 44,3%, o macro-F1 0,474 e o F1 ponderado 0,493. A classe majoritária CAT5 (120 casos) obteve precisão 1,000, mas revocação 0,275; CAT9 e CAT12 chegaram a F1 de 0,857 e 0,840. A assimetria confirma impacto do desbalanceamento e sobreposição semântica, e reforça que a classificação é apoio à decisão, não resposta autônoma. O relatório completo por classe está no Apêndice A.

No PoP, 80,65% dos usuários rastreados que fizeram login assumiram um tíquete; entre os que visualizaram tíquetes, 73,53% assumiram ao menos um. Todos os que assumiram salvaram progresso e 92% finalizaram casos, com média de 6,2 salvamentos por usuário. Oito estagiários finalizaram 114 tíquetes (14,25 por usuário), enquanto 15 especialistas finalizaram 33 (2,2 por usuário). Essa diferença é compatível com triagem volumosa na primeira camada e revisão seletiva na segunda. A utilidade percebida ficou entre 4,51 e 4,66; a facilidade, entre 3,83 e 4,11; e a intenção comportamental, entre 4,03 e 4,34. Facilidade foi o maior preditor da intenção de uso ($\beta = 0,496$), acima da utilidade ($\beta = 0,346$), com 58,5% da variância explicada.

4.3. Engajamento, aprendizagem e experiência de uso

Na primeira avaliação educacional houve atividade em 40 dias, pico de 33 usuários e 2.854 eventos em 12 de janeiro. Vinte e sete usuários concluíram ao menos uma lição e submeteram um quiz; 24 concluíram alguma trilha e 18 algum exercício de programação. A conclusão decresceu de 23 usuários na Trilha 1 para 15 na Trilha 4, padrão compatível com uma sequência progressiva. Três das quatro trilhas tiveram aprovação superior a 80%. A exceção foi a trilha de classificação: 94 tentativas, média 65,27 e aprovação de 54,26%, sinalizando dificuldade conceitual ou desalinhamento entre conteúdo e avaliação. A trilha de *playbooks* DDoS teve média 83,80 e aprovação de 82,86%.

Os elementos gamificados foram usados de fato: registraram-se 468 ganhos de XP por 27 usuários, 107 logins diários por 28, 25 progressões de nível por 16 e 26 compras na loja por dez usuários. Houve ainda 28 tentativas de compra sem saldo,

concentradas em quatro participantes, o que sugere recalibrar a economia. Esses resultados são coerentes com literatura que recomenda analisar a relação entre mecânicas, perfil do usuário e comportamento, em vez de tratar gamificação como efeito uniforme [dos Santos et al. 2025, Meng et al. 2024].

Na segunda avaliação, a média bruta passou de 6,79 acertos no pré-teste ($N = 14$) para 9,17 no pós-teste ($N = 6$), equivalentes a 67,9% e 91,7%. No subconjunto pareado ($N = 5$), o ganho médio foi de três questões, com três correções de erro para acerto e doze resoluções de incerteza para acerto. O resultado é promissor, mas não constitui evidência causal devido à amostra reduzida, ausência de controle e possível familiaridade com o instrumento.

Os questionários do Hackers do Bem indicaram expectativa de desempenho entre 4,23 e 4,54, esforço entre 3,97 e 4,77, condições facilitadoras entre 4,06 e 4,94 e intenção comportamental entre 4,06 e 4,63. Em contraste, a telemetria do MVPv1 registrou 2.448 *dead clicks* e 145 *rage clicks* em 471 visualizações, além da ausência de eventos relacionados à matrícula em trilhas, à visualização de itens e ao fluxo de navegação entre páginas. Os nove e-mails recebidos continham 47 menções a oportunidades de melhoria, sendo 18 relacionadas a bugs e 15 à usabilidade, e 32 avaliações positivas, principalmente sobre o material didático (15) e a estrutura de ensino (10). Em conjunto, os resultados indicam valor percebido e uso efetivo, mas também evidenciam dificuldades de navegação e limitações na observabilidade. Na avaliação de junho, a proporção de *rage clicks* caiu de 0,89% para 0,38%, enquanto a de *dead clicks* aumentou de 15,0% para 28,7%. A análise por proporções e trajetórias evita interpretar a redução do volume absoluto como evidência isolada de melhoria.

4.4. Limitações

O conjunto de incidentes concentra-se em poucos parceiros e categorias, limitando generalização. A qualidade do AutoClass-LFI depende da classe, do serviço de IA e do contexto recuperado; *playbooks* não receberam ainda uma métrica semântica independente. No estudo com residentes, a atividade foi induzida e concentrada, alguns eventos estavam ausentes ou contaminados por desenvolvimento, e questionários mediram sobretudo aceitação, não aprendizagem objetiva. A segunda avaliação teve atrito entre pré e pós-teste. Os próximos ciclos devem ampliar e governar a base, balancear classes, medir correções humanas, registrar funis completos, usar rubricas por competência e conduzir estudos longitudinais com grupos comparáveis.

5. Demonstração e Disponibilidade

5.1. Roteiro da demonstração fim-a-fim

A demonstração será executada como uma narrativa única, preservando o mesmo identificador de caso entre os módulos:

1. **Anonimização:** importação de um pequeno conjunto de incidentes sintéticos ou previamente autorizados; escolha dos campos; detecção de entidades; comparação controlada entre entrada e saída; e inspeção das estatísticas do ANON-LFI.
2. **Categorização e *playbook*:** envio dos registros anonimizados ao AutoClass-LFI; recuperação de casos similares; sugestão de categoria e subcategoria; e geração de um *playbook* em Markdown e Ansible.

3. **Validação manual:** abertura do caso no Sistema PoP, revisão por perfil de estagiário, correção de rótulos ou ações e aprovação final por perfil de especialista, incluindo o histórico das alterações.
4. **Aprendizagem do sistema:** incorporação da decisão validada à base de conhecimento e nova consulta para mostrar como exemplos curados apoiam casos subsequentes, sem afirmar atualização autônoma irrestrita do modelo.
5. **Criação da trilha:** seleção do incidente validado no Hub de Trilhas, definição de competência, geração de lição e quiz, publicação e execução pelo perfil de estudante, com atualização de XP e progresso.

Esse roteiro torna visíveis tanto o caminho dos dados quanto as fronteiras de responsabilidade entre IA, estagiário, especialista e instrutor. A execução principal será feita em navegador e terá duração estimada de 15 minutos. Serão necessários um notebook com navegador atual, Docker e conexão à Internet, um monitor ou projetor e credenciais de demonstração para os três perfis. Um conjunto local reduzido e serviços previamente preparados permitirão concluir o roteiro mesmo com conectividade instável.

5.2. Artefatos e reprodução

O repositório público agregador, disponível em https://github.com/rsmiani/gt-lfi-rnp, documenta a arquitetura e integra os quatro projetos: `anon-lfi`, `lfi-autoclass`, `sistema-pop` e `hackers-do-bem`. Disponibilizado sob licença BSD de 3 cláusulas, o repositório fixa as versões dos componentes por hashes de *commit* e fornece instruções de inicialização, variáveis de ambiente, dados de demonstração, matriz de versões e roteiro de validação, favorecendo a reprodutibilidade e a sustentabilidade dos artefatos. O vídeo da demonstração fim a fim está disponível em https://youtu.be/8F-j4NiD8BI. A reprodução utiliza casos sintéticos e arquivos de exemplo sem identificadores pessoais, uma vez que os dados reais não serão distribuídos.

6. Considerações Finais

O GT-LFI[1] integra privacidade, automação assistida, validação humana e aprendizagem gamificada em um ciclo único. O principal diferencial é a continuidade entre o incidente operacional e o artefato pedagógico: registros reais são preparados, transformados em recomendações provisórias, curados por pessoas e reutilizados em trilhas práticas. As avaliações mostram engajamento, utilidade percebida e sinais descritivos de aprendizagem, ao mesmo tempo que expõem limitações de cobertura, acurácia, usabilidade e desenho experimental.

Como próximos passos, pretende-se ampliar a diversidade da base, formalizar políticas de governança, melhorar revocação e classificação de classes raras, adotar métricas por competência e conduzir avaliações longitudinais com grupos comparáveis. A disponibilização aberta dos componentes e de uma demonstração reproduzível busca permitir que outras instituições inspecionem, adaptem e avaliem o ecossistema em seus próprios contextos.

[1]O presente trabalho foi realizado com o apoio do Programa Hackers do Bem no Domínio Cibernético, executado pela RNP e SENAI, coordenado pela Softex e financiado pelo MCTI com recursos oriundos da Lei das TICs – Lei nº 8.248, de 23 de outubro de 1991. Os autores também reconhecem o uso de ferramentas de Inteligência Artificial Generativa como apoio na revisão e no aprimoramento do código e do manuscrito.

Referências

Alnajim, A. M., Habib, S., Islam, M., AlRawashdeh, H. S., and Wasim, M. (2023). Exploring cybersecurity education and training techniques: A comprehensive review of traditional, virtual reality, and augmented reality approaches. *Symmetry*, 15(12):2175.

Bandel, C. T., Esteves, J. P. R., Guerra, K. P., Bertholdo, L. M., Kreutz, D., and Miani, R. S. (2025). Anonimização de incidentes de segurança com reidentificação controlada. In *Anais do XXV Simpósio Brasileiro de Cibersegurança*, pages 97–113. SBC.

Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly*, 13(3):319–340.

dos Santos, L. F., Oliveira, W., Corrêa de Lima, A., de Castro Junior, A. A., and Hamari, J. (2025). The effects of gamification on learners' engagement according to their gamification user types. *Technology, Knowledge and Learning*.

Gadotti, A., Rocher, L., Houssiau, F., Creţu, A.-M., and de Montjoye, Y.-A. (2024). Anonymization: The imperfect science of using data while preserving privacy. *Science Advances*, 10(29):eadn7053.

Jesus Filho, S. A. d., Santos, Á. S., Araújo, R. D., and Miani, R. S. (2025). Applying llms to the classification of cybersecurity incident tickets in a few-shot scenario. In *Proceedings of the 24th IEEE International Conference on Machine Learning and Applications*. IEEE.

Meng, C., Zhao, M., Pan, Z., Pan, Q., and Bonk, C. J. (2024). Investigating the impact of gamification components on online learners' engagement. *Smart Learning Environments*, 11(1):47.

Nelson, A., Rekhi, S., Souppaya, M., and Scarfone, K. (2025). Incident response recommendations and considerations for cybersecurity risk management: A csf 2.0 community profile. Technical Report NIST SP 800-61r3, National Institute of Standards and Technology.

PostHog (2026). Posthog: Open-source product analytics. https://posthog.com/docs. Acesso em julho de 2026.

Severo, A. S. P., Lautert, D. P., Kreutz, D., Bertholdo, L. M., Pohlmann, M., and Quincozes, S. E. (2025). Categorização de incidentes de segurança utilizando engenharia de prompts em llms. In *Anais do XXV Simpósio Brasileiro de Cibersegurança*, pages 256–272. SBC.

Venkatesh, V., Morris, M. G., Davis, G. B., and Davis, F. D. (2003). User acceptance of information technology: Toward a unified view. *MIS Quarterly*, 27(3):425–478.

Yermilov, O., Raheja, V., and Chernodub, A. (2023). Privacy- and utility-preserving NLP with anonymized data: A case study of pseudonymization. In *Proceedings of the 3rd Workshop on Trustworthy Natural Language Processing*, pages 232–241. Association for Computational Linguistics.

Zhang, J., Bu, H., Wen, H., Liu, Y., Fei, H., Xi, R., Li, L., Yang, Y., Zhu, H., and Meng, D. (2025). When llms meet cybersecurity: A systematic literature review. *Cybersecurity*, 8(1):55.

A. Resultados detalhados e reprodução

Tabela 1. Implementação e contratos entre os componentes.

Componente	Tecnologias	Contrato e persistência
ANON-LFI	Django/DRF, Presidio, React/Vite	CSV → TXT; HTTP com chave de API; SQLite e volumes locais.
AutoClass-LFI	Next.js/Express, Gemini, Pinecone, Firebase	Índices isolados; processamento, fila `available_pop`, esquema e *lineage*.
Sistema PoP	Next.js, React, Firebase	Filas por papel; decisões em `completed_pop`.
Hackers do Bem	Next.js, Firebase, Gemini, Pinecone, PostHog	Conteúdo, progresso, XP, moedas e telemetria.

Tabela 2. Relatório por classe do AutoClass-LFI.

Classe	Precisão	Revocação	F1	Suporte
CAT1	0,667	1,000	0,800	2
CAT2	0,154	0,500	0,235	4
CAT3	0,173	0,944	0,293	18
CAT5	1,000	0,275	0,431	120
CAT7	0,500	0,250	0,333	4
CAT9	0,857	0,857	0,857	7
CAT10	0,000	0,000	0,000	1
CAT12	1,000	0,724	0,840	29
Macro	0,544	0,569	0,474	185
Ponderada	0,876	0,443	0,493	185

A.1. Checklist resumido de reprodução

Para a avaliação do artefato, recomenda-se: (i) clonar o agregador com os hashes de submódulos nele registrados; (ii) conferir as versões de Docker e dos serviços no arquivo de compatibilidade; (iii) copiar o arquivo de variáveis de exemplo sem inserir segredos no repositório; (iv) iniciar os módulos na ordem ANON-LFI, AutoClass-LFI, Sistema PoP e Hackers do Bem; (v) carregar o conjunto sintético; e (vi) executar o roteiro automatizado de sanidade. Saídas esperadas e capturas de referência estarão versionadas no diretório `demo/`.