UNIVERSIDADE
LUSÓFONA

Cibersegurança associada aos RPAs

Trabalho Final de curso

Entrega final

Professor orientador :José Brás

Nome do Aluno: Rafael Lourenço da Silva
Nome do Aluno: Henrique Manuel Pinheiro da Gama Franco

Direitos de cópia

Resumo

Com o aumento da complexidade e dos ataques informáticos, as empresas procuram métodos mais eficazes de se defenderem, pois os seus métodos tradicionais estão cada vez mais frágeis e suscetíveis a ataques às suas infraestruturas de rede, o que pode causar estragos significativos nas organizações.

Neste trabalho final de curso, examinamos a complexidade dos crescentes ataques informáticos e a inadequação dos métodos utilizados na defesa das infraestruturas de rede, propondo uma abordagem inovadora através de um Processo de Automação Inteligente (IPA). Reconhecendo a relevância da cibersegurança e da automação, especialmente quando utilizadas em conjunto com IA, esta combinação permite uma ferramenta promissora para identificar e mitigar vulnerabilidades críticas.

Neste projeto, exploramos a integração de um LLM (Large Language Model) com RPAs, para melhorarmos a precisão e a velocidade de resposta a incidentes, garantindo uma proteção robusta dos sistemas contra novas vulnerabilidades.

Após uma análise das soluções existentes, com destaque para a solução da IBM, e comparando os custos e a eficiência da solução proposta, chegámos à conclusão de que o projeto considera a viabilidade da implementação, abordando os custos de desenvolvimento e estratégias para a gestão dos riscos e da proteção de dados. É apresentada uma análise teórica das redes e do papel do IPA na segurança, evidenciando o valor desta abordagem na identificação de ameaças complexas.

Além disso, o projeto entra em detalhe sobre o processo de automação em conjunto com um LLM para a obtenção das vulnerabilidades da API oficial da NVD. Este é um sistema que fornece vários alertas e informações sobre novas vulnerabilidades.

A metodologia adotada inclui uma abordagem interdisciplinar, com o objetivo de desenvolver um sistema dinâmico que permita uma defesa proativa e eficiente contra estas novas vulnerabilidades, contribuindo assim para a integridade e segurança dos sistemas organizacionais.

Abstract

With the increasing complexity and rise of cyberattacks, companies are seeking more effective methods of defence, as their traditional approaches are becoming increasingly fragile and susceptible to attacks on their network infrastructures, which can cause significant damage to the organisations.

In this final course project, we examine the growing complexity of cyberattacks and the inadequacy of current methods used to defend network infrastructures, proposing an innovative approach through an Intelligent Process Automation (IPA) system.

Recognising the relevance of cybersecurity and automation—especially when used in conjunction with AI—this combination offers a promising tool for identifying and mitigating critical vulnerabilities.

In this project, we explore the integration of a Large Language Model (LLM) with Robotic Process Automation (RPA) to enhance the accuracy and speed of incident response, ensuring robust protection of systems against emerging vulnerabilities.

After analysing existing solutions, with a focus on IBM's approach, and comparing costs and efficiency with our proposed solution, we concluded that the project considers the feasibility of implementation, addressing development costs and strategies for risk management and data protection. A theoretical analysis of networks and the role of IPA in security is presented, highlighting the value of this approach in identifying complex threats.

Moreover, the project details the automation process combined with an LLM to obtain vulnerabilities from the official NVD API. This system provides various alerts and information about new vulnerabilities.

The adopted methodology includes an interdisciplinary approach, aiming to develop a dynamic system that enables proactive and efficient defence against these new vulnerabilities, thus contributing to the integrity and security of organisational systems.

Índice

Lista de Figuras

Lista de Tabelas

1 Introdução

No mundo atual, onde as tecnologias evoluem de forma muito rápida, o mesmo acontece com as ameaças a estas tecnologias. As organizações enfrentam desafios cada vez mais complexos para defender as suas infraestruturas digitais.
Este trabalho final de curso aborda a junção da cibersegurança com uma automação inteligente, com foco na tecnologia de Automação Inteligente de Processos (IPA), para enfrentar problemas reais relacionados com a extração e manipulação de novas vulnerabilidades.
A relevância do estudo reside no cenário crescente de vulnerabilidades informáticas, algo que pode ser verificado no website *nvd.nist.gov*. Os métodos tradicionais de obtenção de vulnerabilidades revelam-se insuficientes face à quantidade e qualidade da informação disponibilizada. Neste contexto, é importante que surjam soluções inovadoras que ofereçam aos profissionais uma forma rápida e eficaz de consultar a informação dessas vulnerabilidades.
Este trabalho é fundamentado em necessidades práticas identificadas em colaboração com a empresa CGI. O objetivo é integrar ferramentas de automação inteligente para monitorizar continuamente novas vulnerabilidades e reportá-las de forma estruturada e eficiente, contribuindo para a segurança e integridade dos sistemas.
A nossa solução é uma aplicação web que, a partir da API do website do NIST, em conjunto com um LLM, permite realizar uma consulta mais detalhada e eficaz, suportada por uma base de dados (IPA) e apresentada visualmente através de um *dashboard* e de uma tabela.

1.1 Enquadramento

A cibersegurança tem se tornado cada vez mais uma área prioritária devido ao grande crescimento das organizações e infraestruturas organizacionais [1]. Este trabalho insere-se nesse contexto, focando-se na integração de uma IPA que cada vez é mais procurado pelas empresas [2].

A automação inteligente surge como uma ferramenta estratégica com grande capacidade na deteção de vulnerabilidades críticas de forma mais rápida possível.

Este trabalho contribui para a evolução das práticas de cibersegurança, oferecendo uma abordagem pratica e adaptada as necessidades reais das organizações.

1.2 Motivação e Identificação do Problema

O aumento da complexidade dos ataques informáticos e das intrusões maliciosas nas infraestruturas de rede globais representa um desafio crescente para os modelos tradicionais de defesa. A deteção manual, por sua vez, exige uma quantidade significativa de tempo e recursos, destacando a necessidade urgente de soluções automatizadas que possam monitorizar e identificar ameaças de forma eficiente e precisa. Este cenário impulsionou a CGI a propor-nos a procura de uma abordagem inovadora, capaz de automatizar o processo de deteção, garantindo uma resposta mais rápida e eficaz a potenciais ataques (alguns requisitos). Dados recentes, conforme ilustrado na Figura 1, indicam um aumento significativo no interesse por temas relacionados com a cibersegurança e a automação, refletindo a prioridade crescente da

automação inteligente na proteção das infraestruturas digitais.

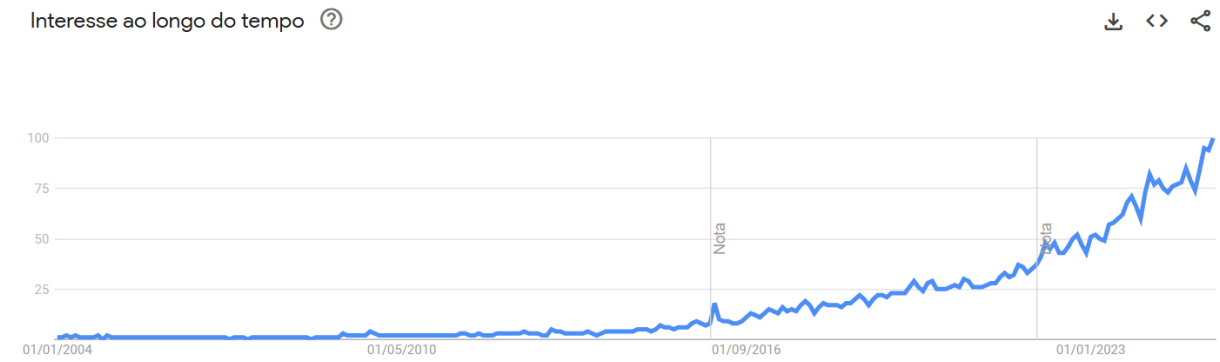

Figura 1-Interesse em cibersegurança desde 2004 (y = Nota de importância 0-100) [3]

Os sistemas de Automação de Processos Robóticos (RPA) em conjunto com o IPA podem potencialmente detetar várias vulnerabilidades nos processos e sistemas de uma organização. Conforme é visível na Figura 2, interesse em procurar informação relacionada com IPA ganhou notoriedade desde 2015 e tem mantido o interesse desde então.

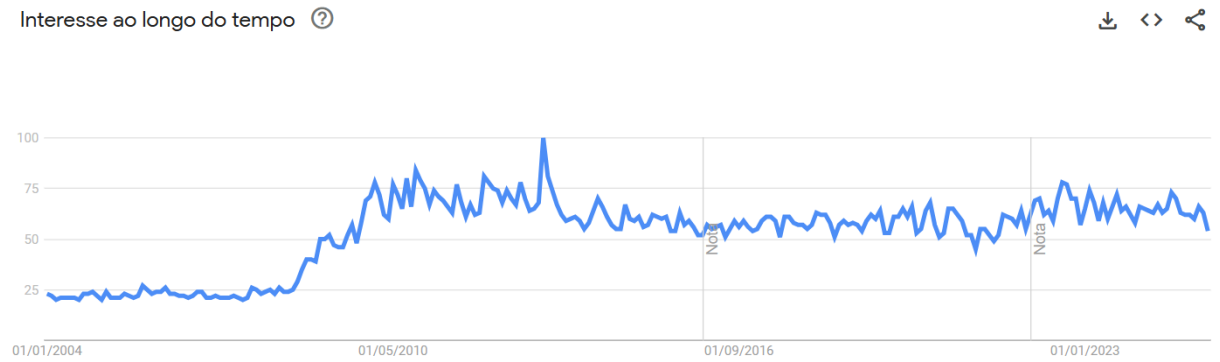

Figura 2-Interesse em IPA desde 2004 (y = Nota de importância 0-100) [4]

Algumas das vulnerabilidades que o IPA pode ajudar a identificar incluem:

Configurações de Segurança Incorretas:

Identificar instâncias em que sistemas, aplicações ou software foram configurados incorretamente, expondo potencialmente dados sensíveis ou tornando o sistema mais suscetível a ameaças informáticas.

Acesso Não Autorizado:

Monitorizar os controlos de acesso e identificar casos em que utilizadores não autorizados tentam aceder a sistemas ou dados críticos, ajudando a prevenir potenciais violações de dados ou manipulação não autorizada de dados.

Fugas de Dados:

Detetar casos em que dados sensíveis estão a ser transmitidos de forma insegura ou fora de redes seguras, ajudando a prevenir fugas de dados e garantindo a conformidade com regulamentos de proteção de dados.

Questões de Conformidade:

Ajudar as organizações a identificar situações de não conformidade com políticas internas ou regulamentos externos, garantindo que todos os processos estejam de acordo com os padrões necessários e reduzindo o risco de penalizações regulamentares.

Cifragem Inadequada:

Identificar casos em que os dados não estão adequadamente cifrados, garantindo que informações sensíveis estejam protegidas contra acesso ou interceção não autorizados.

Ataques de Phishing e Engenharia Social: Ajudar a detetar padrões ou atividades suspeitas que possam indicar a presença de ataques de phishing ou tentativas de engenharia social, permitindo que as organizações adotem medidas preventivas para proteger os seus sistemas e funcionários. Ao monitorizar ativamente essas vulnerabilidades, os sistemas de IPA podem contribuir para a postura de segurança geral de uma organização, ajudando a identificar e abordar proactivamente ameaças potenciais antes que se transformem em incidentes de segurança mais significativos.

O nosso objetivo será principalmente a automação e deteção de vulnerabilidades de Acessos não autorizados e Fugas de dados.

1.3 Objetivos

O principal objetivo deste Trabalho Final de Curso é desenvolver uma solução automatizada que, a partir de um website, permita ao departamento de segurança informática estar constantemente atento a novas vulnerabilidades que possam ser críticas para os seus sistemas. Outro objetivo é possibilitar a obtenção dessas vulnerabilidades de forma rápida e com o menor número de erros possível, garantindo ainda a escalabilidade da ferramenta, permitindo assim obter resultados mais eficazes e em menos tempo.
Pretendemos também utilizar Inteligência Artificial (IA) para a extração e manipulação da informação recebida, de forma a potenciar a eficiência, automatizar processos e obter *insights* relevantes a partir de grandes volumes de dados. Através de técnicas como o processamento de linguagem natural (NLP), *machine learning* e análise preditiva, pretendemos identificar padrões, classificar informações, filtrar conteúdos relevantes e transformar dados brutos em conhecimento útil para apoio à tomada de decisão.

1.4 Estrutura do Documento

Na Secção 1 é apresentada uma breve introdução do trabalho.

Na secção 2 é apresentada a análise da viabilidade e pertinência do trabalho desenvolvido.

Na secção 3 é falado sobre as especificações e modelação do trabalho.

Na secção 4 falamos sobre o Estado da Arte.

Na secção 5 falamos sobre o enquadramento teórico e científico e sobre as tecnologias utilizadas.

Na secção 6 falamos sobre a solução proposta e sobre algum do trabalho já desenvolvido.

2 Pertinência e viabilidade

2.1 Pertinência

Com a evolução dos sistemas tecnológicos a aumentarem cada vez mais, implicando também o aumento da exposição dos dados e dos crimes dentro da cibersegurança é cada vez mais importante termos os nossos dados protegidos da forma mais segura possível.

Estudos indicam que o mercado de cibersegurança relacionada com a automação vai ter um grande aumento como mostra o estudo seguinte:

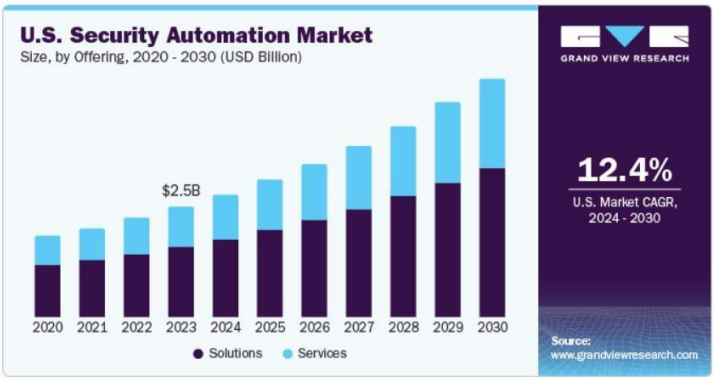

Figura 3-Aumento do mercado de automação na cibersegurança

Automatizar a deteção de Acessos Não Autorizados e Fugas de Dados é crucial por várias razões fundamentais, como por exemplo:

Resposta Imediata: A automação permite uma deteção rápida e contínua de acessos não autorizados e fugas de dados, possibilitando uma resposta imediata para interromper essas atividades maliciosas antes que causem danos significativos.

Figura 4-Resultados do questionário

Redução de Riscos e Danos: A deteção precoce de acessos não autorizados e fugas de dados ajuda a mitigar riscos e reduzir potenciais danos financeiros, reputacionais e legais que podem resultar de violações de segurança.

Conformidade Regulatória: A automação na deteção dessas violações ajuda as organizações a manter a conformidade com as regulamentações de proteção de dados, evitando penalidades e sanções associadas a violações de privacidade e segurança de dados.

Proteção de Dados Sensíveis: Ao automatizar a deteção de fugas de dados, as organizações podem proteger informações sensíveis e confidenciais, garantindo a privacidade e a segurança dos dados dos clientes e funcionários.

Eficiência Operacional: A automação libera recursos humanos valiosos, permitindo que as equipas de segurança de TI se concentrem em atividades de análise mais complexas e na implementação de medidas preventivas mais robustas para fortalecer a postura de segurança geral da organização.

Prevenção de Ataques Cibernéticos: A deteção automatizada de acessos não autorizados e fugas de dados pode ajudar a identificar tentativas de intrusão e atividades maliciosas antes que os hackers tenham a oportunidade de comprometer significativamente os sistemas e dados da organização. A automação desses processos críticos não apenas aumenta a eficácia e a agilidade da deteção, mas também fortalece a capacidade de uma organização de proteger-se contra ameaças informáticas e salvaguardar a integridade dos seus dados confidenciais.

A solução que desenvolvemos e pertinente maioritariamente por diariamente salvar tempo aos utilizadores da WebApp que necessitam de fazer esta recolha de vulnerabilidades de forma "manual" o que custa tempo útil do trabalhador que por sua vez se reflete no custo laboral desse mesmo trabalhador na empresa, enquanto na nossa solução, todo este processo é feito de forma automática e organizada.

2.2 Viabilidade

O desenvolvimento de uma IPA envolve custos que podem variar significativamente, dependendo da escalabilidade da plataforma escolhida e da qualidade do desenvolvimento do processo robótico.
Pode ser utilizada uma IA com maior capacidade de processamento, especialmente em casos em que são necessárias respostas ou análises muito complexas. No entanto, para que isso aconteça localmente, é necessário dispor de uma infraestrutura capaz de suportar estes modelos. Esta abordagem implica um investimento considerável em componentes físicos, mas permite que a empresa mantenha total controlo sobre os dados, uma vez que tudo é processado internamente e, se desejado, em sistemas isolados da internet. Esta configuração é ideal em ambientes que exigem elevados níveis de segurança.
Por outro lado, existe também a possibilidade de utilizar uma API externa, como a do ChatGPT ou Gemini, fornecida por plataformas como a OpenAI, Google ou Cloudflare. Esta abordagem tem a vantagem de não exigir hardware próprio potente, pois todo o processamento é feito na *cloud*. Os custos variam consoante o volume de utilização e são geralmente calculados com base nos *tokens* de entrada e de saída [5].

No entanto, é importante considerar que, ao recorrer a uma API externa, os dados enviados para processamento podem circular por servidores fora da empresa e, em muitos casos, fora do país. Apesar de estas plataformas aplicarem medidas de segurança robustas, existe sempre o risco potencial de fugas de dados ou de acesso não autorizado.
Esses fatores são cruciais, pois uma plataforma escalável permite que o sistema cresça com o aumento da procura, enquanto um desenvolvimento de alta qualidade reduz os riscos de falhas e os custos de manutenção. Quando o desenvolvimento de uma IPA não atinge um padrão adequado de qualidade, as organizações podem enfrentar custos elevados de manutenção, uma vez que problemas de inoperacionalidade nos robôs projetados para proteger e otimizar os seus sistemas acarretam despesas significativas. Estes custos de manutenção, aliados ao tempo e aos recursos necessários para corrigir falhas, comprometem a rentabilidade esperada da automação.
Para que uma implementação de IPA seja bem-sucedida, não basta apenas adotar a tecnologia; é fundamental adotar uma abordagem abrangente e estratégica. Isso inclui a integração das melhores práticas de segurança, para garantir que a automação não introduza vulnerabilidades e que os dados críticos da empresa permaneçam protegidos. Além disso, a gestão cuidadosa do ciclo de vida dos produtos de IPA é essencial, desde o planeamento e desenvolvimento, passando por testes rigorosos, até à manutenção e eventual atualização.
A implementação de IPA deve ser vista como um compromisso contínuo, em que a organização monitoriza e aprimora o desempenho do robô, ajustando-o conforme necessário para que continue a agregar valor e a responder às necessidades operacionais, garantindo, assim, a sustentabilidade da automação ao longo do tempo.
Além disso, visto que se trata de uma tecnologia recente e em constante evolução e descoberta, à medida que vai progredindo, poderá até ser possível automatizar outras tarefas relacionadas com as vulnerabilidades, permitindo assim um progresso contínuo deste projeto.

2.3 Análise Comparativa com Soluções Existentes

2.3.1 Soluções existentes

Dentro do contexto deste trabalho e aprimorando ainda mais a eficiência e precisão dos sistemas de segurança já existentes pretende-se fazer uma Automatização Avançada de Resposta a

Incidentes: Integrar sistemas de resposta a incidentes totalmente automatizados que possam isolar instantaneamente partes afetadas da rede, interromper o acesso não autorizado e minimizar o impacto de eventuais fugas de dados. Para além disso, pretende-se fazer uma Análise Comportamental em Tempo Real: Desenvolver sistemas de análise comportamental em tempo real que possam detetar anomalias instantaneamente e tomar medidas corretivas imediatas para prevenir violações de segurança.

Como podemos ver na figura 3 o preço mensal seria 2704.63 para 10 robôs enquanto para um recurso em início de carreira seria um salário de 1400 € brutos que dá um total de cerca de 2000€ de custo para uma empresa hipotética. Conforme descrito na página da IBM [6].

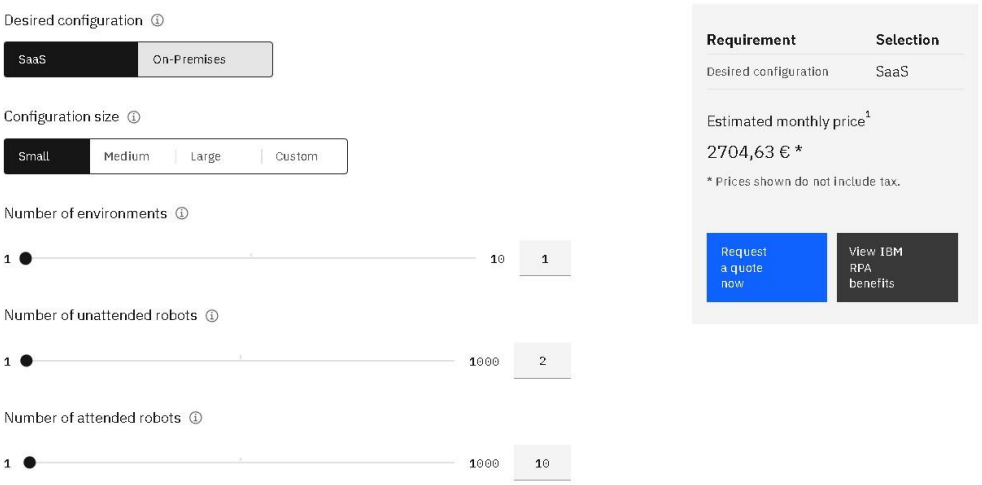

Figura 5-IBM Robotic Process Automation Price Estimator

2.3.2 Análise de benchmarking

Funcionalidades	IBM	McAfee
Resposta a incidentes em tempo real	x	
Automatização de Resposta a Ameaças	x	x
Análise de Comportamento Anômalo	x	
Segurança de Dados e Prevenção contra Perda de Dados (DLP)	x	x
Gestão de Patches	x	

Resposta a Incidentes em Tempo Real*	x	x
Soluções de Segurança para Blockchain		x

Tabela 1-Funcionalidades

A nossa solução propõe uma automação avançada que otimiza significativamente a resposta e deteção de incidentes, ao contrário de outras abordagens no mercado que não oferecem esta solução. Esta automação elimina processos manuais demorados e reduz a possibilidade de falhas humanas na obtenção e na comunicação das vulnerabilidades. Alem disso ao utilizar a nossa proposta a equipa de segurança consegue focar-se em outras atividades mais estratégicas.

2.4 Proposta de inovação e mais-valias

A solução apresenta uma abordagem inovadora no que toca a gestão de vulnerabilidades, automatizando a extração de dados, geração de relatórios e monitorização continua diferenciando-se por sua proatividade, personalização e custo-eficiência.

Para a empresa a nossa abordagem promete sustentabilidade maximizando o retorno ao reduzir os custos operacionais, mas ainda é uma alternativa acessível às soluções comerciais como também provoca um alívio na dependência de processos manuais.

2.5 Identificação de oportunidade de negócio

A solução pode ser explorada comercialmente como uma plataforma SaaS acessível, direcionada a pequenas e médias empresas. O diferencial está na sua acessibilidade, personalização de relatórios e monitorização contínua, oferecendo uma alternativa eficaz e de baixo custo às soluções comerciais tradicionais. Este modelo fomenta o empreendedorismo tecnológico, atendendo à crescente procura por cibersegurança eficiente e adaptada às necessidades do mercado.

3 Especificação e Modelação

3.1 Analise de Requisitos

Requisitos funcionais:

Identificação	Descrição	MoSCoW
Conexão com a API do NVD	O sistema deve conectar-se com a API do NVD para conseguir extrair as vulnerabilidades	Must Have
Autenticação e Autorização na API	Utilizar a key para fazer a autenticação e as requisições à API	Must Have
Extração de Dados	Extrair os dados relevantes	Must Have

Processamento da Descrição	Preparar a descrição para a análise pelo LLM	Must Have
Integração com LLM para Categorização	Utilizar o LLM para que, a partir da descrição, atribua uma ou várias categorias a cada vulnerabilidade	Should Have
Validação e Limpeza dos Dados	Utilizar formas automáticas para validar e limpar os dados extraídos da API	Must Have
Armazenamento Estruturado	Inserir os dados na base de dados	Must Have
Associação entre Tabelas	Garantir que as tabelas estão todas relacionadas e que os dados foram inseridos corretamente	Must Have
Geração de Queries via LLM	Utilizar o LLM para transformar perguntas em queries de SQL para fazer consultas à base de dados	Must Have
Interface Web para Consulta:	Desenvolver um website para que os utilizadores possam fazer perguntas ao LLM	Must Have
Exibição de Resultados	Apresentar as respostas das consultas de forma clara e estruturada	Must Have
Suporte a Múltiplos Filtros e Parâmetros	Permitir que os utilizadores refinem as suas consultas utilizando filtros adicionais	Should Have

Tabela 2-Requisitos Funcionais

Requisitos não funcionais:

Identificação	Descrição	Área
Performance e tempo de resposta	O sistema deve responder às consultas e processamentos em poucos segundos, garantindo uma experiência fluida ao utilizador	Performance
Escalabilidade	A arquitetura deve suportar o aumento no volume de dados e números de vulnerabilidades sem perder desempanho	Performance e Escalabilidade
Proteção de dados	Implementar mecanismos para evitar o acesso não autorizado e proteger os dados armazenados contra-ataques	Segurança
Confiabilidade e Disponibilidade	O sistema deve garantir alta disponibilidade com estratégias de redundância e backup para minimizar downtime	Modelo Operacional
Modularidade do Código	Estruturar o código de forma modular, facilitando manutenções e futuras extensões do sistema	Implementação e Desenvolvimento

Documentação Completa	Manter uma documentação atualizada e detalhada dos módulos de integração, processamento, armazenamento e interface	Implementação e Desenvolvimento
Testabilidade	Incluir testes unitários, de integração e de performance para assegurar a estabilidade do sistema	Testes
Usabilidade e Interface Intuitiva	Desenvolver uma interface web de fácil uso, mesmo para usuários com baixo conhecimento técnico	Personalização
Escalabilidade	O sistema deve utilizar padrões abertos (REST, JSON, SQL) para facilitar a integração com outros serviços ou APIs	Integração
Manutenibilidade	Garantir que o sistema seja fácil de atualizar e corrigir, com uma arquitetura que facilite a identificação de problemas	Manutenção
Flexibilidade para Atualizações	Permitir a fácil atualização dos componentes do sistema (como o LLM ou integrações com novas APIs) sem interromper o funcionamento do sistema	Implementação e Desenvolvimento

Tabela 3-Requisitos não funcionais

3.2 Use case

O diagrama de casos de uso da WebApp descreve, de forma concisa, as interações entre quatro atores (Utilizador,API NVD, Agendador 24 h e CloudFlare LLM) e os principais serviços do sistema. O utilizador pode autenticar-se, definir preferências, consultar vulnerabilidades, pesquisar por texto ou CVE ID, filtrar por categoria e fazer consultas em linguagem natural. Internamente, a extração periódica via API NVD dispara processos de limpeza, importação de dados históricos e armazenamento numa BD relacional; as consultas acionam a categorização automática por LLM e a geração de queries SQL. O Agendador garante a atualização diária, enquanto o LLM provê tanto a classificação inteligente das vulnerabilidades

quanto a interpretação das perguntas em NL, garantindo dados sempre atualizados e facilmente acessíveis.

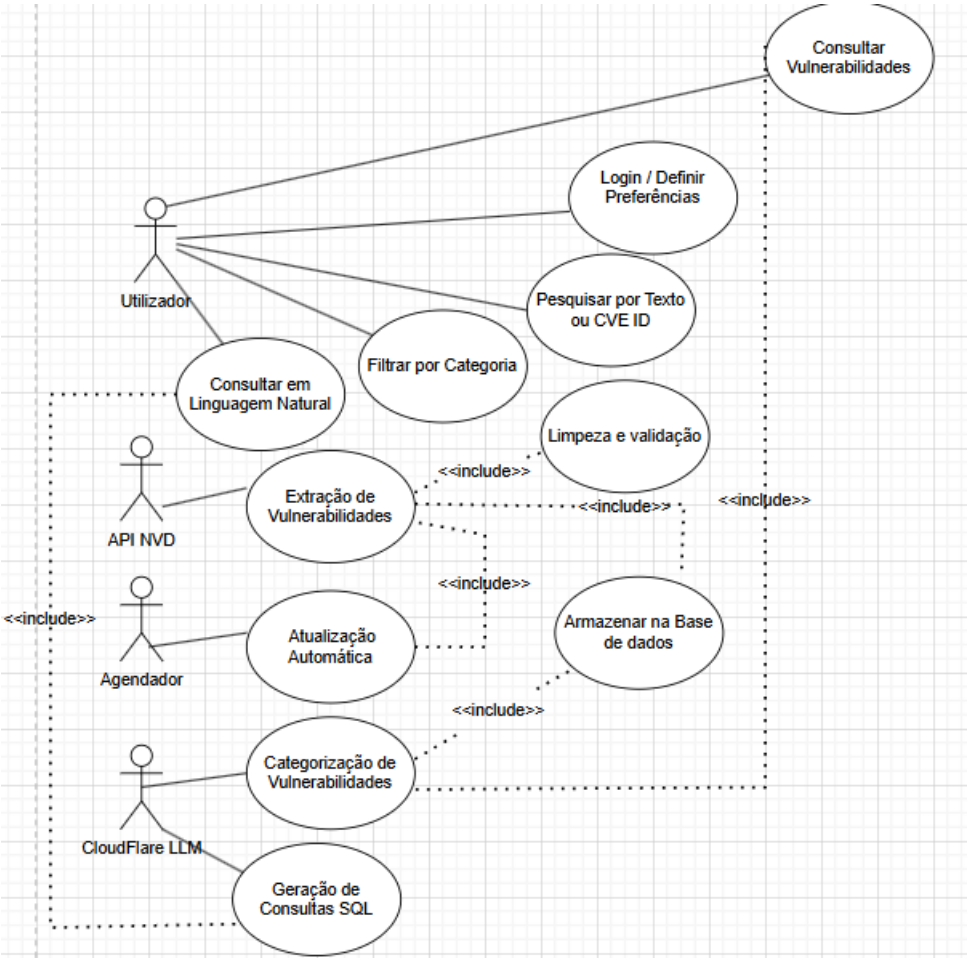

Figura 6-Use case

4 Estado da arte

Com o aumento das ameaças cibernéticas, cresce também a adoção de soluções automatizadas para deteção de acessos não autorizados e fugas de dados. Organizações em todo o mundo recorrem tanto a plataformas comerciais líderes, como o IBM QRadar e o McAfee Total Protection for DLP, como a soluções desenvolvidas internamente, adaptadas às suas necessidades específicas.
A solução apresentada neste contexto — um sistema automatizado de deteção, categorização e armazenamento de vulnerabilidades, baseado em dados do NVD (*National Vulnerability Database*) — oferece uma abordagem alternativa, focada na automação total do ciclo de ingestão de vulnerabilidades. Utilizando Python, bases de dados SQLite e modelos de linguagem como o LLaMA, esta ferramenta distingue-se pelas seguintes capacidades:

- Integração automática com a NVD: coleta contínua e personalizada de vulnerabilidades mais recentes.

- Classificação automática via IA: categorização dinâmica das vulnerabilidades com uso de modelos de linguagem.
- Base de dados própria e consultável: permite análises específicas com geração automatizada de queries em SQL.
- Enriquecimento com múltiplas fontes: consolida e armazena referências externas para cada vulnerabilidade.

Em comparação, soluções comerciais como IBM QRadar e McAfee oferecem funcionalidades mais amplas de SIEM (Security Information and Event Management) e DLP (Data Loss Prevention), incluindo:

- Análise comportamental avançada (IBM QRadar) [14]: deteção em tempo real de atividades anómalas com base em correlação de eventos.
- Prevenção ativa de perda de dados (McAfee DLP) [15] : aplicação de políticas de segurança sobre tráfego de dados em endpoints e servidores.
- Integração com ambientes corporativos complexos: suportam múltiplos protocolos, sistemas operacionais e ambientes cloud/híbridos.
- Interface gráfica e dashboards em tempo real: foco em visualização e orquestração de incidentes.

Apesar dessas vantagens, tais plataformas podem apresentar custos elevados de aquisição, licenciamento e manutenção. Por outro lado, a solução proposta neste projeto representa uma opção leve, escalável e de código aberto, ideal para:

- Ambientes de investigação e desenvolvimento;
- Pequenas e médias empresas com equipas técnicas internas;
- Casos em que se deseje personalizar ao máximo a deteção e categorização de vulnerabilidades, sem depender de plataformas proprietárias.

Conclui-se que, enquanto as soluções comerciais oferecem pacotes integrados e prontos para produção em larga escala, a abordagem baseada em automação personalizada proposta neste projeto destaca-se pela flexibilidade, controlo total dos dados e possibilidade de extensão contínua a baixo custo.

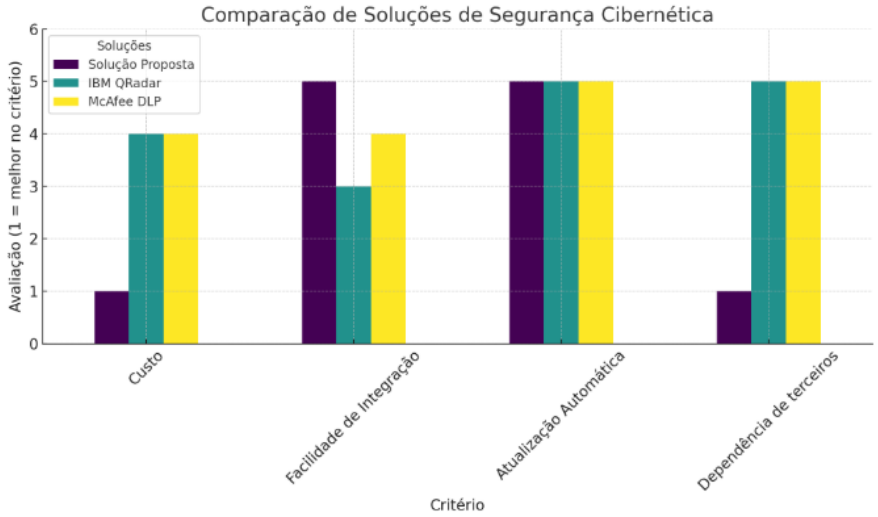

Figura 7-Comparação de soluções existentes

4.1 Análise de custos e viabilidade:

Item	Estimativa de Custo	Observações
Desenvolvimento Inicial	€1 000 – €6 000	Varia conforme complexidade e nível de integração com sistemas legados.
Infraestrutura	€0 – €300/mês	Servidor cloud básico ou VPS; praticamente nulo se usar infra open-source já existente.
Manutenção & Suporte	€500 – €2 000/mês	Inclui updates de modelo, correções de bugs e pequenas adaptações.
LLM Local (on-premise)	+ €500 – €10 000/mês	Custo de servidores GPU dedicados; maior controlo e privacidade.
LLM via API (pay-per-use)	€0,02 – €0,10 por 1 000 tokens	Custo variável conforme volume de consultas; sem capex de hardware.
SaaS (e.g. ChatGPT)	€20 – €500/mês por utilizador	Modelo "tudo pronto"; licenciamento por usuário; SLA e suporte incluídos.

Tabela 4-Análise de custos

4.1.1 Retorno sobre Investimento (ROI)

- **Faixa Geral de Economia**: 30 % a 90 % de redução de custos em comparação a soluções comerciais fechadas.
- **Principais Fontes de Economia**:
  - Corte de licenças proprietárias (até 90 % em cenários enxutos).[14]
  - Automatização de tarefas repetitivas e redução de retrabalho manual [15].
  - Visibilidade e insights em tempo real, evitando desperdícios operacionais.

4.1.2 Vantagens e Desvantagens das Opções de Implantação

1. **LLM Local (On-Premise / Private Cloud)**
   a. **Vantagens**:
      i. Total controle sobre dados e compliance.
      ii. Latência interna otimizada.
   b. **Desvantagens**:
      i. Alto investimento inicial em hardware (servidores GPU).[16]
      ii. Necessidade de equipe especializada para manutenção.
2. **LLM via API (Provedores Hospedados)**
   a. **Vantagens**:
      i. "Pay-as-you-go": flexibilidade orçamentária conforme uso.[17]
      ii. Rápido deployment, sem capex em infraestrutura.[18]
   b. **Desvantagens**:
      i. Custos variáveis podem ficar altos em volumes intensivos.
      ii. Dependência externa; cuidado com privacidade de dados.
3. **SaaS (ChatGPT, Anthropic, etc.)**

a. **Vantagens**:
    i. Interface e backend gerenciados pelo fornecedor.[19]
    ii. Upgrades, segurança e suporte incluídos no plano.
b. **Desvantagens**:
    i. Licenciamento por usuário; menos personalização do modelo.[20]
    ii. Potencial lock-in e restrições de expansão.

5 Enquadramento teórico e científico do problema

Cibersegurança:

No decorrer dos anos, os sistemas informáticos evoluíram de meros meios de comunicação para infraestruturas computacionais ubíquas. As redes tornaram-se maiores, mais complexas, mais rápidas e altamente dinâmicas. O uso diário e generalizado das tecnologias de computação e redes em todos os aspetos da vida transformou as questões de segurança informática em temas críticos para as organizações e, inclusivamente, em questões de segurança nacional [1]. Ou seja, os sistemas precisam de ser projetados e testados com a segurança em mente desde o início, e não apenas como um acréscimo ou uma reflexão posterior.
Ao projetar um sistema seguro, a segurança não deve ser apenas mais um atributo desejável do sistema. É necessário que o foco não esteja apenas numa solução amigável para os seus utilizadores e eficiente, mas sim na existência de uma solução equilibrada em todos os aspetos. [7].

Automação de Processos Robóticos (RPA):

A Automação de Processos Robóticos (RPA) é uma tecnologia que permite automatizar tarefas repetitivas e baseadas em regras, normalmente realizadas por humanos, utilizando robôs de software ou inteligência artificial. Esses robôs são programados para executar tarefas com precisão e consistência, seguindo instruções definidas por desenvolvedores, que podem usar métodos como gravação de ecrã, definição de variáveis e fluxos de trabalho específicos. Algumas tarefas que o RPA pode realizar incluem aceder a aplicações, copiar e colar dados, enviar emails, preencher formulários, gerar relatórios periódicos, entre outros.
De acordo com Van der Aalst, "RPA é um termo abrangente para ferramentas que operam na interface do utilizador de outros sistemas de computador"[8]. Embora formas tradicionais de automação de processos, como gravação de ecrã, scraping e macros, também dependam da interface do utilizador, a principal característica do RPA é a sua capacidade de identificar elementos da interface diretamente, em vez de confiar em coordenadas de ecrã ou seleções XPath. Isso permite uma interação mais robusta e menos propensa a erros, melhorando a fiabilidade do processo.
Desde 2020, os fornecedores de RPA relatam um aumento significativo na procura pelas suas soluções, com projeções que indicam um crescimento ainda maior até 2030. Além do setor empresarial, as ferramentas de RPA eso a ser cada vez mais aplicadas em áreas como auditoria,forense digital e automação industrial, mostrando-se essenciais para a transformação digital desses setores.

Figura 8-Aumento do mercado RPA [2]

Com o avanço da indústria, o RPA tornou-se uma ferramenta essencial para otimizar tarefas operacionais e oferecer vantagens competitivas, especialmente em setores que procuram a digitalização de processos. Através da integração de dados obtidos de dispositivos conectados, o RPA permite que as empresas automatizem uma grande variedade de tarefas comerciais de rotina, melhorando a eficiência e reduzindo o tempo necessário para as executar. Diferentemente dos métodos tradicionais de automação, o RPA opera sobre a infraestrutura de TI existente, não necessitando de alterações significativas nos sistemas subjacentes. Este fator reduz a complexidade da implementação, evita altos níveis de intrusão nos sistemas corporativos e permite uma integração rápida e de baixo custo. Estudos conduzidos pela Deloitte apontam que a implementação de RPA traz melhorias significativas em aspetos essenciais para as empresas: [9]

- **92% de melhoria na conformidade**, o que garante que os processos estejam alinhados com normas e regulamentações.
- **86% de aumento na produtividade**, permitindo que as empresas realizem tarefas mais rapidamente e com maior eficiência.
- **90% de melhoria na qualidade**, reduzindo erros e aumentando a confiabilidade das operações.
- **59% de redução de custos**, aliviando as pressões financeiras e liberando recursos para outras iniciativas estratégicas.

No entanto, por ser uma tecnologia relativamente nova, o RPA ainda apresenta desafios no que diz respeito à segurança e à gestão de riscos. Como qualquer aplicação de software, o RPA está sujeito a vulnerabilidades. Para reduzir esses riscos, é fundamental que as empresas sigam princípios de segurança específicos e integrem o RPA em um quadro robusto de governança [10].

O primeiro passo para garantir a segurança é estabelecer uma estrutura de governança apropriada. Isso inclui a criação de uma estratégia de avaliação de riscos, que identifique e analise potenciais ameaças e vulnerabilidades associadas ao uso do RPA. Além disso, é essencial

implementar controles de segurança, monitoramento contínuo e políticas de atualização, assegurando que os robôs operem em conformidade com os padrões de segurança da empresa e possam responder de forma ágil a novos riscos emergentes.

Automação de processos inteligente (IPA):

Um IPA representa a evolução das tecnologias de automação. Esta tecnologia combina RPA com Inteligência Artificial (IA) e *Machine Learning* (ML). Enquanto o RPA tradicional se limita a executar tarefas repetitivas e baseadas em regras pré-programadas, o IPA eleva a automação para o próximo nível, permitindo que os sistemas aprendam e melhorem continuamente à medida que realizam mais tarefas.
Com a integração de IA e ML, os IPA são capazes de interpretar e processar grandes volumes de dados de forma rápida e eficiente. Além disso, conseguem identificar padrões, tomar decisões complexas e até mesmo adaptar-se a novas situações sem intervenção humana. Ao contrário de uma automação estática, um IPA tem um enorme potencial de aprendizagem contínua, o que o torna mais eficiente e dinâmico a longo prazo.
Prevê-se um grande aumento na procura de soluções relacionadas com esta tecnologia, conforme ilustrado no gráfico a seguir.

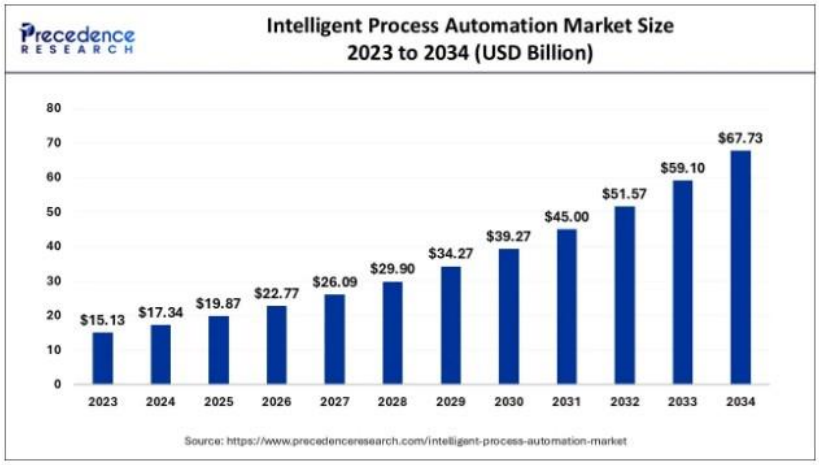

Figura 9- Aumento do IPA [11]

LLM - Large Language Model:

Um LLM, ou *Large Language Model*, é um modelo de inteligência artificial que utiliza técnicas de *Machine Learning* para compreender e gerar linguagem humana.
Os LLMs baseiam-se em redes neurais e adotam técnicas de processamento de linguagem natural (NLP) para fornecer respostas, utilizando algoritmos de aprendizagem profunda (*Deep Learning*) e arquiteturas com *transformers* para processar, compreender e gerar linguagem natural. Estes modelos são treinados com grandes volumes de dados, o que lhes permite captar semânticas contextuais, tornando-os capazes de executar diversas tarefas, desde a geração de texto e traduções automáticas, entre muitas outras.

Estes modelos convertem o texto em representações numéricas que são processadas de forma mais eficiente e inteligente. Em vez de utilizar uma simples tabela numérica que atribui um número a cada palavra, os LLMs utilizam vetores multidimensionais (*word embeddings*). Estes vetores são capazes de captar semânticas, o que permite que palavras com significados ou contextos semelhantes fiquem próximas no espaço vetorial.
A utilização de *word embeddings* e *transformers* permite superar grandes limitações dos modelos anteriores de representação, que não conseguiam captar relações semânticas entre as palavras. A partir desta abordagem, é possível o reconhecimento de contextos, a geração de linguagem natural e a aprendizagem em larga escala.

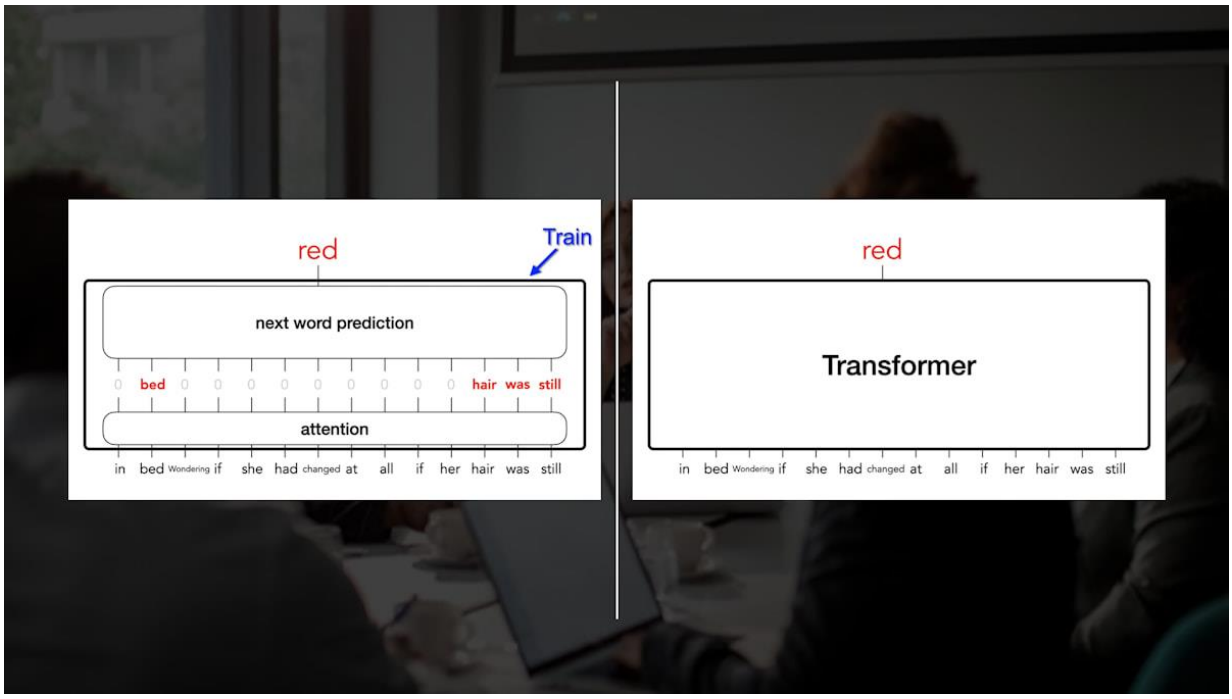

Figura 10-Funcionamento de um LLM - [12]

A utilização dos LLMs em contextos empresariais tem tido um crescimento significativo nos últimos anos, impulsionando assim a necessidade de soluções mais eficientes e inteligentes em diversas áreas do negócio.
Este aumento é devido a necessidade de automatizar tarefas repetitivas desde assistentes virtuais, automação de tarefas, criação de conteúdo, analise de dados, entre outros. Os LLMs permitem as empresas adaptar cada modelo as suas necessidades especificas pois permitem o seu treinamento com dados específicos para cada tarefa [13].
NLP – Processamento de Linguagem Natural
O NLP (Natural Language Processing) é uma área da inteligência artificial e ciência da computação que visa capacitar as máquinas para interpretar, analisar e gerar linguagem humana de maneira eficaz. Utiliza algoritmos de aprendizado de máquina e modelos estatísticos para transformar texto não estruturado em dados úteis e interpretáveis, facilitando a automação de processos, a extração de informação relevante e a compreensão do contexto presente em grandes volumes de dados textuais.

6 Trabalho anterior à solução proposta

1. Contexto e Objetivo

A CGI fez-nos uma proposta para implementar uma solução automatizada de coleta e análise de vulnerabilidades listadas pelo NIST (National Institute of Standards and Technology). O principal objetivo era disponibilizar, em tempo real, um conjunto estruturado de dados sobre vulnerabilidades (CVE IDs, descrições, datas, pontuações CVSS, referências etc.) para suportar decisões de segurança da informação.

2. Web Scraping Automatizado

Nesta fase implementamos uma solução utilizando selenium(Python) que obtia automaticamente as vulnerabilidades presentes no website NIST e guardavas num Excel. Após isso, a partir de um LLM local tentamos treiná-lo para identificar as categorias das vulnerabilidades a partir da descrição. Devido à falta de processamento dos nossos computadores chegamos á conclusão de que esta solução não seria exequível.

3. Integração via API Oficial NIST

Apos a solução previa fizemos alguma investigação e descobrimos que o website do NIST tinha uma API publica que permitia realizar um pedido GET das vulnerabilidades e informações associadas.

4. Arquitetura de Armazenamento e Escalabilidade

Para armazenarmos a informação utilizamos uma base de dados relacional, o que possibilita uma maior escalabilidade, pesquisa e organização dos dados em comparação a um ficheiro excel.

5. LLM- CloudFlare

Visto que o LLM localmente não tinha a performance desejada encontramos a solução de utilizar o LLM de forma gratuita a partir da API da empresa CloudFlare com alguns limites diários.

6. Interface gráfica em Flas

Inicialmente por questões praticas decidimos criar uma WebApp em flask para permitir ao utilizador consultar a informação de forma mais organizada mesmo que de forma minimalista (framework).

7 Solução Proposta

A WebApp desenvolvida tem como principal objetivo facilitar a monitorização e análise de vulnerabilidades de segurança informática, com base em dados provenientes da API oficial do NIST. As suas principais funcionalidades incluem:

- **Consulta de Vulnerabilidades:** Apresentação das vulnerabilidades mais recentes obtidas através da API do NIST, com dados como ID (CVE), descrição, data de publicação, pontuação de gravidade (CVSS) e referências.

- **Classificação Automática:** O sistema utiliza um modelo de linguagem (LLM) para analisar a descrição de cada vulnerabilidade e atribuir-lhe automaticamente uma categoria relevante, como Windows, Linux, Redes, Web, entre outras.
- **Pesquisa por Texto ou CVE ID:** Os utilizadores podem procurar vulnerabilidades específicas introduzindo o ID ou termos relevantes, permitindo localizar rapidamente casos de interesse.
- **Filtros por Categoria:** A interface inclui opções para filtrar os resultados por categoria atribuída, facilitando a análise contextual.
- **Armazenamento Persistente:** Toda a informação é armazenada numa base de dados relacional, o que permite histórico, análises e escalabilidade da aplicação.
- **Atualização Automática:** Processo automático de 24 em 24 horas (alterável, se necessário) que garante que a base de dados está sempre atualizada.
- **Consultas com Linguagem Natural:** Os utilizadores podem fazer perguntas diretamente em linguagem natural (por exemplo, "Quais são as vulnerabilidades críticas de junho?"), com o LLM a interpretar a intenção da pergunta e a devolver os resultados corretos utilizando a categorização previamente indicada.
- **Sistema de Login:** A solução contém um sistema de login que permite aos utilizadores com conta registada escolher diferentes categorias de preferência, o que afeta tanto a tabela como o dashboard, embora na tabela os filtros possam ser limpos para mostrar todas as vulnerabilidades.

A imagem abaixo representa a arquitetura conceptual da solução proposta, ilustrando as principais componentes e as suas interações. O diagrama destaca o fluxo de dados desde a extração das vulnerabilidades através da API oficial do NVD, passando pela limpeza e categorização com o Motor IPA e LLM, até à disponibilização e consulta dos dados pelos utilizadores através da interface web, com destaque para o papel do LLM nas pesquisas em linguagem natural e categorização inteligente.

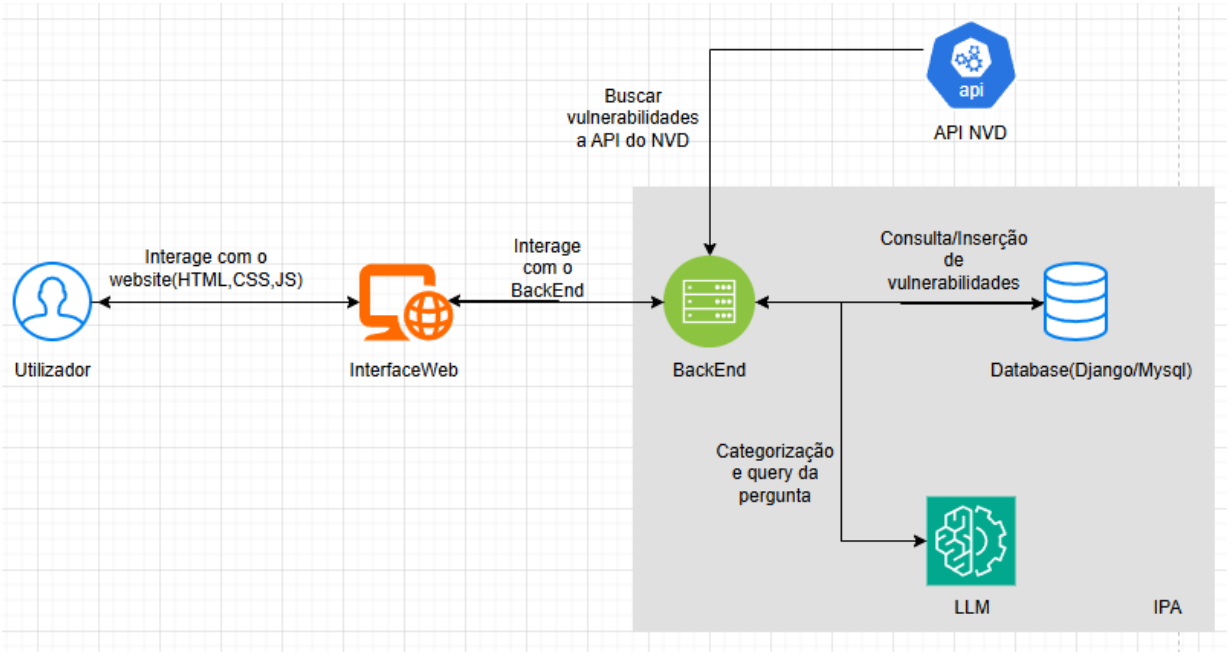

Figura 11-Arquitetura da solução

InterfaceWeb

7.1 Base de dados

Desenvolvemos uma base de dados relacional utilizando models do Django, estruturada em quatro tabelas principais, concebidas para armazenar apenas os dados essenciais sobre cada vulnerabilidade de forma a otimizar a performance e a escalabilidade do sistema ao contrário da solução anterior.

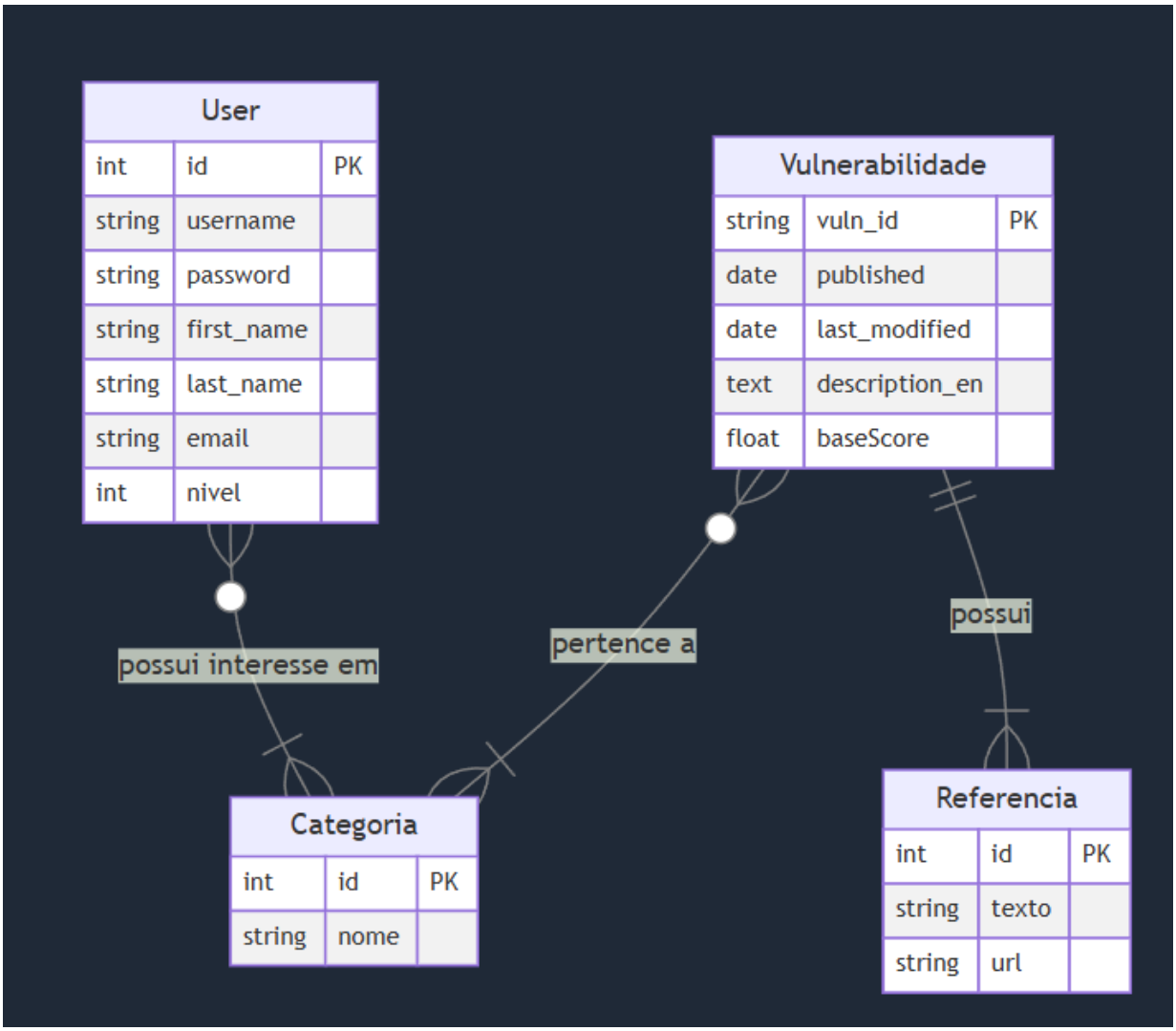

Figura 12-Estrutura da base de dados

7.2 Extração das vulnerabilidades

Para garantirmos a extração dos dados de forma eficiente e com alta fiabilidade, decidimos utilizar a API oficial do NIST onde recebemos a informação a partir de um JSON, abandonando o WebScraping. Esta escolha pode ser comparada à diferença entre tentar tirar água de um poço com um balde preso por uma corda, ou simplesmente abrir uma torneira ligada diretamente à fonte.

Processo de extração:

- **Conexão com a API:** Utilizamos uma interface de comunicação com a API do NIST, que permite extrair os dados estruturados das vulnerabilidades
- **Limpeza e validação:** Antes de inserir os dados na base de dados, um conjunto de algoritmos realiza a limpeza de cada vulnerabilidade, removendo caracteres especiais, corrigindo formatações erradas e garantindo a integridade dos dados. Esta etapa é fundamental para evitar erros na consulta e processamento futuro.
- **Extração de dados históricos:** A API do NIST disponibiliza dados desde antes do ano 2000, possibilitando a expansão do histórico caso exista alguma necessidade de análise a dados mais antigos.
- **Atualizações automáticas:** O sistema está configurado para verificar periodicamente a API e atualizar a base de dados apenas com as novas vulnerabilidades que não estejam registadas, assegurando que a informação esteja sempre atualizada.

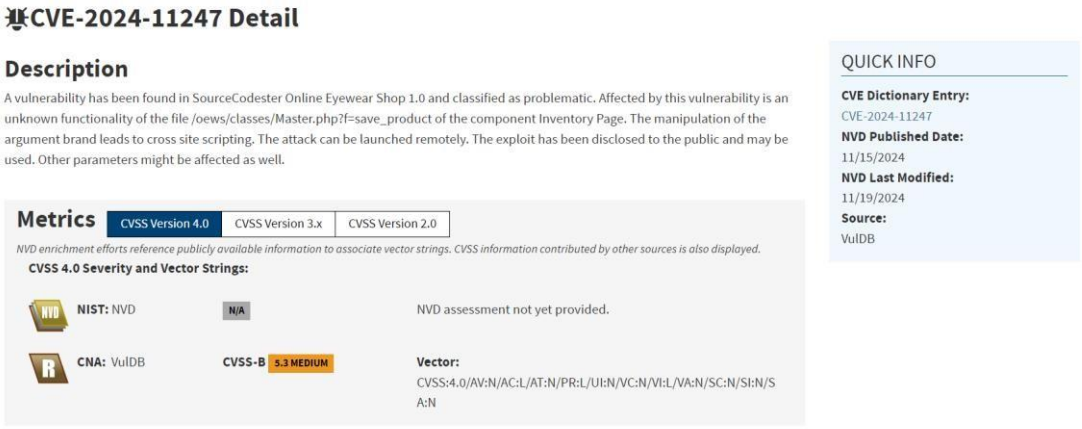

Figura 13-Informações sobre as vulnerabilidades

7.3 Utilização do LLM na categorização e geração de consultas

O uso de um modelo de linguagem (LLM), especificamente o llama-3-8b-instruct, é uma componente crucial desta solução, atuando tanto na categorização das vulnerabilidades quanto na geração de queries dinâmicas para consultas.

7.3.1 Categorização de Vulnerabilidades

Após a extração e limpeza dos dados, o LLM analisa as descrições das vulnerabilidades para atribuir uma ou mais categorias pertinentes, como Windows, SQL, Google, entre muitas outras. Este processo envolve uma análise e uma atribuição das categorias de forma totalmente automática.

Este modelo está a ser executado num servidor da empresa CloudFlare devido a falta de capacidade de processamento que temos disponível localmente. Utilizamos este LLM a partir de uma API providenciada pela CloudFlare de forma gratuita.

7.3.2 Geração de consultas automatizadas

O mesmo LLM é utilizado para interpretar consultas formuladas em linguagem natural e traduzi-las em queries SQL que possam ser executadas na base de dados do django. Esta funcionalidade permite que os utilizadores interajam com o Sistema de forma intuitiva e preservando os dados da base de dados tendo em conta que o LLM não tem acesso direto á mesma.

Antes da query ser executada garantimos que não existe nenhuma tentativa de alteração ou eliminação da base de dados nem acesso a tabela dos utilizadores.

7.4 Interface Web

A interface web foi desenvolvida a pensar na simplicidade e na utilidade prática para equipas técnicas que precisam de consultar e analisar vulnerabilidades de forma rápida e eficiente.

Quando o utilizador entra no website, tem acesso a uma página inicial que explica de forma detalhada as funcionalidades da nossa WebApp. Esta página serve como ponto de partida para que qualquer pessoa, mesmo sem conhecimento prévio da aplicação, possa perceber o que o sistema faz, como funciona a recolha de dados, e quais as ferramentas disponíveis para análise de vulnerabilidades.

Depois dessa introdução, o utilizador pode aceder à área principal da aplicação, onde se encontra uma dashboard. Esta dashboard foi desenhada para oferecer uma visão clara e imediata do estado atual das vulnerabilidades registadas na base de dados, com foco nas categorias mais relevantes para cada utilizador.

A dashboard inclui os seguintes componentes:

- **Número total de vulnerabilidades** atualmente registadas na base de dados.
- **Top 5 vulnerabilidades mais críticas** dos últimos 3 meses (com pontuação base superior a 9).
- **Número de vulnerabilidades críticas** publicadas nos últimos 3 meses.
- **Últimas 5 vulnerabilidades** associadas às categorias selecionadas pelo utilizador. Caso o utilizador não tenha nenhuma categoria selecionada mostra as 5 categorias mais recentes.
- **Top 5 categorias mais afetadas**, com base no volume de vulnerabilidades.
- **Gráfico mensal**, que mostra a evolução do número de vulnerabilidades ao longo do último ano.

Logo após esta visualização inicial, o utilizador pode aceder à tabela detalhada das vulnerabilidades. Esta tabela apresenta informações como o ID do CVE, descrição da falha, data de publicação, pontuação CVSS e links para referências externas. Um aspeto fundamental da aplicação é o uso de **categorias automáticas**. Caso o utilizador tenha o login feito e tenha selecionado vulnerabilidades do seu interesse, ao entrar nesta página a tabela vai mostrar automaticamente apenas as vulnerabilidades que estão relacionadas com as categorias em questão. Caso não tenha nenhuma categoria selecionada, ainda é possível filtrar a tabela manualmente, escolhendo as categorias pretendidas.

Nesta tabela também é possível realizar pesquisas através de uma **barra de pesquisa** que se encontra no topo da página. O utilizador pode escrever palavras-chave ou perguntas completas, como por exemplo "vulnerabilidades críticas de maio", e o sistema interpreta a intenção com recurso ao modelo de linguagem (LLM), devolvendo os resultados filtrados diretamente na própria tabela. Esta

funcionalidade combina a flexibilidade da linguagem natural com a estrutura de visualização tradicional, o que torna a experiência bastante intuitiva.

A aplicação também disponibiliza uma barra de pesquisa intuitiva onde se pode procurar vulnerabilidades por palavras-chave ou diretamente pelo ID (como "2025-1234"). A resposta é rápida e pensada para oferecer uma experiência fluida, mesmo em dispositivos com menos capacidade.

A interface foi desenhada com foco na simplicidade, clareza e rapidez. Evitámos sobrecargas visuais, mantendo tudo direto ao ponto, com tempos de carregamento reduzidos. O objetivo é que qualquer profissional técnico, mesmo sem especialização em segurança, consiga utilizar a aplicação de forma natural e produtiva.

A interface foi desenhada com foco na simplicidade, clareza e rapidez. Evitámos sobrecargas visuais, mantendo tudo direto ao ponto, com tempos de carregamento reduzidos. O objetivo é que qualquer profissional técnico, mesmo sem especialização em segurança, consiga utilizar a aplicação de forma natural e produtiva.

7.5 Abrangência

Nesta solução as disciplinas que achamos pertinentes para este TFC foram:

- Fundamentos de programação, Linguagens de programação 1, Algoritmia e estrutura de dados, Linguagens de programação 2 – Estas disciplinas foram importantes para aprendermos a programar e a produzir código de qualidade.
- Engenharia de software – Engenharia de software foi importante para nos conseguirmos organizar e conseguir entregar o trabalho a tempo.
- Base de dados – Foi fundamental para conseguirmos aceder automaticamente a uma base de dados, como também configurá-la.
- Inteligência Artificial – Foi crucial para a utilização e implementação de um LLM.
- Programação Web – Foi utilizada para a criação do website.
- Data Science – Foi importante para conseguirmos manipular toda a informação.

7.6 Trabalho Futuro

Para o trabalho futuro é imperativo implementar a introdução de vulnerabilidades para cada utilizador, ou seja, cada utilizador poder ter uma pequena base de dados com as suas próprias vulnerabilidades fazendo com que possam ver não só o histórico da introdução das vulnerabilidades como também a informação de cada vulnerabilidade. Para além disso o Dashboard pode e deve ser melhorado para se tornar mais interativo e informativo sobre as vulnerabilidades de cada utilizador com gráficos representativos das mesmas. Outro aspeto importante que pode ser implementado é a obtenção de vulnerabilidades a partir de outras fontes de informação sem ser o NIST. Para fazer esta implementação apenas é necessário preparar e ajustar a informação no momento da extração para depois ser colocada na base de dados.

8 Calendário

Figura 14-Calendário de Gaant 1

Figura 15-Calendário de Gaant 2

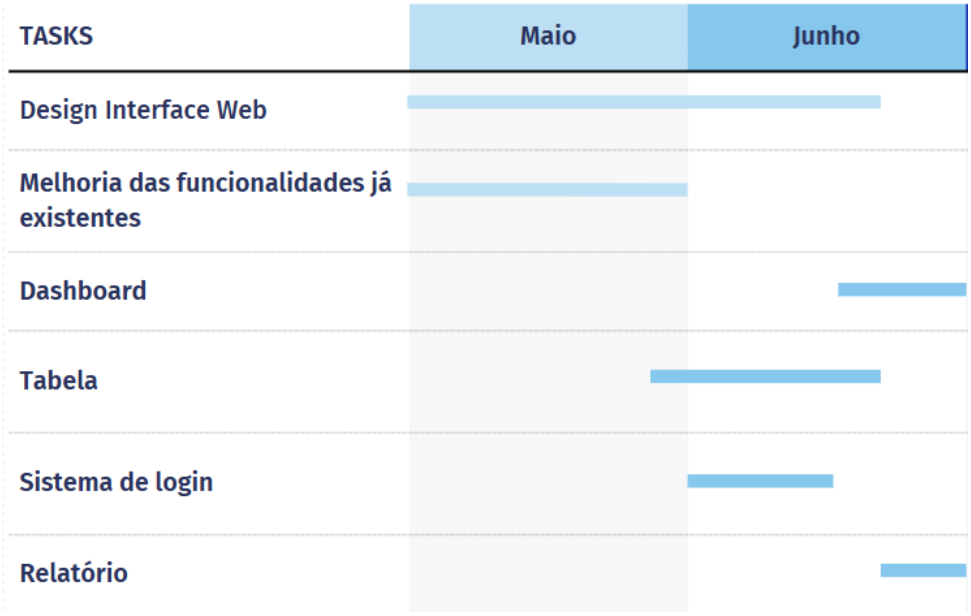

Figura 16-Calendário de Gaant 3

Bibliografia

1. Vodafone Portugal alvo de ciberataque – VodafonePortugal.https://www.vodafone.pt/press-releases/2022/2/vodafonehttps://www.vodafone.pt/press-releases/2022/2/vodafone-portugal-alvo-de-ciberataque.htmlportugal-alvo-de-ciberataque.html

2. "Increase of robotic process automation: https://www.statista.com/statistics/1259903/robotic-processhttps://www.statista.com/statistics/1259903/robotic-process-automation-market-size-worldwide/automation-market-size-worldwide/

3. "Cibersegurança - Explorar - Google Trends.": https://trends.google.com/trends/explore?date=all&q=Ciberseguran%C3%A7a& hl=pt-PT

4. "Robotic Process Automation - Explorar - Google Trends.": https://trends.google.com/trends/explore?date=all&q=Robotic%20Process%20Automation&hl=pt-PT

5. LLM Pricing Calculator: https://www.helicone.ai/llm-cost

6. "Pricing - IBM Robotic Process Automation.": https://www.ibm.com/products/robotic-process-automation/pricing

7. R. A. Kemmerer, "Cybersecurity," Proceedings - International Conference on Software Engineering, pp. 705–715, 2003, doi: 10.1109/ICSE.2003.1201257.

8. W. M. P. van der Aalst, M. Bichler, and A. Heinzl, "Robotic Process Automation," Business and Information Systems Engineering, vol. 60, no. 4, pp. 269–272, Aug.2018, doi: 10.1007/S12599-018-0542-4/FIGURES/1.

9. "Melhorias da RPA nas empresas"-Deloitte. https://www2.deloitte.com/content/dam/Deloitte/us/Documents/process-and-operations/us-global-rpa-survey-infographic.pdf

10. R. A. Kemmerer, "Cybersecurity," Proceedings - International Conference on Software Engineering, pp. 705–715, 2003, doi: 10.1109/ICSE.2003.1201257.

11. "Increase of Intelligence Process Automation": https://www.precedenceresearch.com/intelligent-process-automation-market

12. Como funciona um LLM: https://pplware.sapo.pt/inteligencia-artificial/o-que-e-e-como-funciona-um-large-language-model-llm/

13. The evolution of LLM Adoption in Industry Data Curation Practices: https://arxiv.org/abs/2412.16089

14. https://pmc.ncbi.nlm.nih.gov/articles/PMC7480774/?utm_source=chatgpt.com
15. https://willdom.com/blog/cost-savings-with-ai-automation/?utm_source=chatgpt.com
16. https://lumenalta.com/insights/understanding-the-cost-to-setup-an-ai-data-center-updated-2025?utm_source=chatgpt.com
17. https://www.withorb.com/blog/pricing-ai-agents?utm_source=chatgpt.com
18. https://www.pwc.com/us/en/services/consulting/business-transformation/data-analytics.html?utm_source=chatgpt.com
19. https://www.pwc.com/us/en/tech-effect/ai-analytics/ai-predictions.html
20. https://www.gartner.com/en/information-technology/glossary/software-as-a-service-saas
21. https://www.ibm.com/qradar

Anexo 1 – Questionário

https://docs.google.com/forms/d/1wa-9D3zWGWtbb2qpODVwG-X9_NACW49A-G649NhQUBE/edit

Anexo 2 – Vídeo da solução implementada

https://youtu.be/kytHo2KGm1E

Glossário

LEI	Licenciatura em Engenharia Informática
LIG	Licenciatura em Informática de Gestão
TFC	Trabalho Final de Curso
RPA	Automação robótica de processos
IPA	Automação Inteligente de Processos