1. Introdução

Este relatório tem como objetivo analisar os riscos à proteção de dados pessoais associados à aplicação que realiza a raspagem de informações do Sofascore e armazena os dados em arquivos CSV para posterior utilização por um modelo de inteligência artificial (IA), baseado no Ollama com DeepSeek, a fim de responder às perguntas dos usuários. A análise segue a metodologia STRIDE e a modelagem de ameaças de Torr (2005), permitindo a identificação de vulnerabilidades e a proposição de medidas de mitigação.

A arquitetura da aplicação envolve um sistema de raspagem que coleta dados públicos do Sofascore, armazena essas informações localmente e as disponibiliza para consulta via IA. Essa abordagem traz desafios relacionados à segurança dos dados processados, armazenamento e controle de acesso, sendo necessário avaliar os riscos envolvidos e definir estratégias para mitigação.

2. Riscos Identificados no Sistema

Para identificar e priorizar os riscos associados à aplicação, utilizamos uma abordagem integrada que combina a modelagem de ameaças proposta por Torr (2005) com a estrutura STRIDE. Essa combinação possibilita a análise sistemática das vulnerabilidades, levando em conta a probabilidade de ocorrência e o impacto potencial de cada ameaça. A seguir, detalhamos as principais categorias de riscos identificadas:

2.1 Falsificação de Identidade (Spoofing)

Aplicando os princípios de Torr (2005), constatamos que a ausência de um mecanismo robusto de autenticação expõe o sistema a ataques de spoofing. Nesse cenário, um invasor pode se passar por um usuário legítimo para obter acesso indevido aos dados raspados do Sofascore. Essa vulnerabilidade compromete a confiabilidade dos dados utilizados pela inteligência artificial, impactando diretamente a veracidade das respostas fornecidas. A análise ressalta a necessidade de implementar controles de autenticação (por exemplo, uso de tokens ou autenticação multifator) para reduzir essa ameaça.

2.2 Adulteração de Dados (Tampering)

A integridade dos dados é um pilar fundamental para o funcionamento da aplicação. É crucial identificar pontos onde a manipulação possa ocorrer, uma vez que a transmissão e o armazenamento dos dados em arquivos CSV abrem margem para ataques de tampering. Nesses casos, um agente mal-intencionado pode alterar as informações coletadas, resultando em respostas incorretas pela IA e comprometendo a confiabilidade do serviço. Mecanismos como hashing, assinaturas digitais e controle de versões são recomendados para mitigar esse risco.

2.3 Repúdio (Repudiation)

A falta de um sistema de logs e registros detalhados dificulta a rastreabilidade das ações realizadas dentro da aplicação. Sem a capacidade de auditar e reconstruir o histórico de operações, um usuário mal-intencionado pode negar a autoria de ações indevidas, ampliando o risco de repúdio. Dessa forma, é fundamental registrar todas as interações de forma segura para assegurar a responsabilidade e viabilizar investigações posteriores, mantendo a integridade do controle de acesso.

2.4 Exposição de Informações Confidenciais (Information Disclosure)

Mesmo que os dados coletados sejam originalmente públicos, a agregação e o processamento podem revelar informações sensíveis, como padrões de acesso, metadados e comportamentos dos usuários. Essa exposição pode servir como porta de entrada para investigações aprofundadas ou para ataques direcionados. Portanto, a mitigação dessa ameaça envolve a aplicação de criptografia aos dados, o armazenamento seguro e a minimização das informações sensíveis mantidas no sistema.

2.5 Negação de Serviço (Denial of Service - DoS)

A continuidade dos serviços oferecidos pela aplicação depende da sua disponibilidade, que pode ser comprometida por ataques de DoS. Esses ataques exploram falhas na infraestrutura ou sobrecarregam o sistema com um volume excessivo de requisições, inviabilizando tanto o processo de raspagem quanto a consulta via IA. Medidas como rate limiting, balanceamento de carga e monitoramento constante do tráfego são essenciais para manter a resiliência da aplicação frente a esse tipo de ameaça.

2.6 Elevação de Privilégio (Elevation of Privilege)

Quando os controles de acesso não são suficientemente rigorosos, há o risco de elevação de privilégio, em que um atacante explora vulnerabilidades para obter permissões superiores às necessárias. Essa situação compromete não apenas os dados, mas toda a integridade do sistema. Adotar o princípio do menor privilégio, limitando o acesso de cada componente ao estritamente necessário para sua função, aliado a auditorias regulares e revisões dos mecanismos de controle, é fundamental para mitigar esse risco.

3. Estratégias de Mitigação dos Riscos

Para combater essa ameaça, a aplicação adota mecanismos de autenticação robustos, integrando autenticação multifator e a utilização de tokens de acesso, como os JSON Web Tokens (JWT). Essa estratégia garante que somente usuários devidamente verificados possam interagir com o sistema, permitindo, ainda, a verificação rigorosa da origem das requisições e o monitoramento contínuo para detectar e bloquear tentativas de acesso indevido, conforme as recomendações de Torr (2005).

3.1 Falsificação de Identidade (Spoofing)

Para combater essa ameaça, a aplicação adota mecanismos de autenticação robustos, integrando autenticação multifator e a utilização de tokens de acesso, como os JSON Web Tokens (JWT). Essa estratégia garante que somente usuários devidamente verificados possam interagir com o sistema, permitindo, ainda, a verificação rigorosa da origem das requisições e o monitoramento contínuo para detectar e bloquear tentativas de acesso indevido, conforme as recomendações de Torr (2005).

3.2 Adulteração de Dados (Tampering)

A integridade dos dados é assegurada por meio da proteção das comunicações através de protocolos seguros, como o HTTPS, e da implementação de técnicas de hashing e assinaturas digitais para validar os arquivos armazenados. Além disso, o controle de versões aliado a auditorias regulares permite a identificação e reversão de quaisquer alterações não autorizadas, mantendo a confiabilidade dos dados processados pela aplicação.

3.3 Repúdio (Repudiation)

A mitigação do repúdio é realizada por meio da implementação de um sistema de registros detalhados, que documenta todas as interações e transações no sistema, registrando data, hora, identificação dos usuários e a natureza das operações realizadas. Esses registros, protegidos por assinaturas digitais, asseguram a rastreabilidade das ações e facilitam a responsabilização, alinhando-se aos princípios estabelecidos por Torr (2005).

3.4 Exposição de Informações Confidenciais (Information Disclosure)

Mesmo considerando que os dados coletados são originalmente públicos, a aplicação adota a criptografia forte tanto para os dados armazenados quanto para os dados em trânsito, prevenindo a exposição de informações sensíveis. Essa medida, aliada à prática de minimização de dados, reduz significativamente a superfície de ataque e protege metadados que possam comprometer a privacidade dos usuários.

3.5 Negação de Serviço (Denial of Service - DoS)

Para garantir a disponibilidade contínua da aplicação, são implementadas estratégias como o controle do volume de requisições por meio de técnicas de rate limiting e o balanceamento de carga entre servidores. Além disso, o monitoramento constante do tráfego e a utilização de mecanismos de defesa, como firewalls de aplicação, contribuem para a identificação e mitigação de padrões anormais de acesso, prevenindo ataques de DoS e DDoS.

3.6 Elevação de Privilégio (Elevation of Privilege)

A aplicação adota o princípio do menor privilégio, assegurando que cada componente e usuário possua apenas as permissões estritamente necessárias para suas funções. Essa estratégia é reforçada por auditorias periódicas dos controles de acesso e pela utilização de

técnicas de isolamento, como namespaces e cgroups, que limitam a possibilidade de exploração de vulnerabilidades que possam permitir a obtenção de privilégios indevidos.

4. Conclusão

A análise de impacto à proteção de dados pessoais identificou riscos significativos na aplicação que realiza raspagem de dados do Sofascore e os disponibiliza para consulta via IA. A implementação das medidas propostas é essencial para reduzir vulnerabilidades, garantindo a integridade, segurança e conformidade do sistema com boas práticas de proteção de dados.

Monitoramento contínuo, revisão periódica dos mecanismos de segurança e adaptação a novas ameaças devem ser priorizados para assegurar a proteção das informações processadas pela aplicação.