Desafio: Análise Financeira

Edital: CPSI

1. O Desafio (Solução Inovadora)

A solução será uma plataforma unificada de inteligência financeira, integrando e correlacionando automaticamente dados de múltiplas fontes para construir uma visão 360º do investigado. Ela proverá a reconstrução da linha do tempo patrimonial, assegurando auditabilidade completa e compartimentação rigorosa de informações sigilosas.

2. Objetivo

Disponibilizar solução tecnológica integrada destinada à análise financeira investigativa, capaz de correlacionar e analisar dados bancários, fiscais, societários e patrimoniais provenientes de múltiplas fontes heterogêneas, inclusive Relatórios de Inteligência Financeira (RIFs) do COAF, com geração de inteligência investigativa qualificada e suporte à adoção de medidas judiciais, observada, de forma rigorosa, a compartimentação dos dados sigilosos por Caso investigativo.

A solução deverá permitir a realização de análises financeiras de forma simples, eficiente e acessível, inclusive por usuários sem perfil técnico avançado, promovendo celeridade investigativa, aumento da efetividade das medidas assecuratórias e fortalecimento da sustentabilidade jurídica das análises produzidas.

Para tanto, a solução deverá contemplar, de forma integrada, os seguintes eixos funcionais:

a) Ingestão e integração de dados Capacidade de ingestão automatizada, tratamento e padronização de dados bancários, fiscais, financeiros, societários, patrimoniais e cadastrais, oriundos de arquivos eletrônicos, consultas web, bases públicas, sistemas conveniados e integrações via webservices/APIs, com controle de acesso e compartimentação por Caso investigativo;

b) Inteligência analítica e correlação avançada Consolidação automática dos registros e correlação entre pessoas físicas e jurídicas, bens, movimentações financeiras e vínculos indiretos, com suporte à identificação de patrimônio oculto, interposição de pessoas, estruturas societárias complexas e trilhas financeiras relevantes. Deverá incorporar mecanismos de Inteligência Artificial, Processamento de Linguagem Natural e suporte a LLMs, incluindo arquitetura RAG segmentada por Caso, com respostas fundamentadas exclusivamente nos dados do respectivo acervo e indicação das fontes utilizadas;

c) Visualização, relatórios e suporte decisório Disponibilização de dashboards dinâmicos, visualizações analíticas integradas, grafos de relacionamento e recursos de business intelligence, bem como a emissão de relatórios técnicos customizáveis, com rastreabilidade das fontes, integridade dos dados e suporte direto à fundamentação de medidas cautelares patrimoniais;

d) Gestão do conhecimento e inteligência institucional Constituição de acervo institucional estruturado, seguro e auditável, com capacidade de armazenamento, organização,

compartilhamento e reutilização de documentos, dados e análises financeiras, viabilizando sua aplicação em investigações correlatas e promovendo a consolidação de inteligência institucional. A solução deverá permitir, ainda, a ingestão estruturada de dados e documentos diretamente pelos usuários responsáveis pelas investigações, em ambiente unificado.

Busca-se, com isso, elevar o padrão técnico das Análises Financeiras, reduzir significativamente o tempo médio de produção dessas análises, ampliar a capacidade institucional de identificação e constrição de ativos de origem ilícita e aumentar a efetividade das medidas patrimoniais requeridas pelo Ministério Público.

3. Oportunidades de Mercado

1. Centralidade da Recuperação de Ativos na Persecução Penal.

A recuperação de ativos é dimensão essencial do combate à corrupção e ao crime organizado. Privar o agente criminoso do proveito econômico constitui a medida mais eficaz de prevenção e recomposição do patrimônio público, amparada pelos arts. 91 e 91-A do Código Penal e legislações especiais (Leis n°s 12.850/13 e 9.613/98). A efetividade destas medidas exige análise financeira rápida e integrada, hoje prejudicada por consultas manuais a sistemas desconectados, elevando o risco de dissipação patrimonial.

2. Inadequação das Ferramentas Atuais.

O modelo vigente apresenta limitações que comprometem a eficiência investigativa:

Fragmentação e Retrabalho: Consultas individuais em múltiplas interfaces sem consolidação automática.

Invisibilidade de Vínculos: Dificuldade em detectar correlações em estruturas societárias complexas ou interposição de pessoas (laranjas).

Gargalo em Dados Não Estruturados: Processamento manual de RIFs, PDFs e planilhas, limitando a escala da análise.

3. Potencial de Replicabilidade e Impacto.

A arquitetura modular permitirá a expansão para outros ramos do Ministério Público e instituições do sistema de justiça (Polícias, CGU e Tribunais de Contas). A solução deve atuar como um hub de inteligência institucional, permitindo a reutilização de trilhas investigativas e padrões de lavagem de dinheiro, transformando o conhecimento individual em acervo estratégico compartilhado.

4. Diferenciais de Inovação Esperados.

A solução deve transcender as ferramentas tradicionais de BI, integrando:

Ingestão Inteligente e Automatizada: Integração direta com ecossistemas de dados bancários (SIMBA) e fiscais, com tratamento especializado de formatos heterogêneos.

Análise Assistida por IA (RAG): Consulta ao acervo do caso em linguagem natural, com respostas fundamentadas e rastreabilidade total da fonte, eliminando alucinações.

Segregação Nativa por Caso: Compartimentação lógica e vetorial rigorosa, garantindo que dados sigilosos de uma investigação jamais contaminem ou sejam acessíveis por outras.

Detecção de Padrões e Grafos: Identificação automática de anomalias financeiras e vínculos indiretos, revelando estruturas de ocultamento de bens.

BI Investigativo e Relatórios Cautelares: Dashboards dinâmicos e geração de relatórios técnicos robustos para instruir pedidos judiciais de constrição patrimonial.

Colaboração e Ingestão Ativa: Capacidade de o usuário alimentar o ambiente com documentos específicos da investigação, promovendo a centralização do conhecimento.

4. Requisitos Funcionais

ID	Descrição	PoC	Homologação
1	Central de Casos: A tela inicial do usuário autenticado deve exibir os Casos aos quais tem acesso, com indicadores visuais de status (ativo, arquivado, aguardando dados), data da última atividade, pendências de análise e acesso rápido às funcionalidades mais utilizadas.	5	15
2	Fluxo Guiado de Criação de Caso: O sistema deve conduzir o usuário por um assistente passo a passo para criação de novo Caso, com preenchimento de dados básicos, definição de usuários autorizados, configuração de fontes de dados iniciais (como SIMBA ou RIFs) e confirmação de compartimentação, com validação em tempo real.	5	15
3	Notificações e Alertas: O sistema deve exibir notificações	5	15

ID	Descrição	PoC	Homologação
	contextuais ao usuário sobre eventos relevantes do Caso, como conclusão de ingestão de dados, alertas de padrões atípicos detectados, documentos pendentes de revisão e falhas de extração, sem necessidade de o analista consultar cada módulo individualmente.		
4	Exportação de Visualizações: A interface deve permitir a exportação de grafos e dashboards em formatos compatíveis para apresentações (ex: PDF, PNG).	-	5
5	Anotações e Marcações Contextuais: O usuário deve poder inserir anotações textuais vinculadas a registros específicos (transações, pessoas, documentos etc.), com registro de autoria e data/hora, visíveis apenas aos usuários autorizados no Caso e integradas ao acervo vetorial para consulta via RAG.	10	15
6	Linha do Tempo de Eventos: Geração automática de um registro cronológico e interativo de ações, transações financeiras suspeitas ou de grande vulto, alterações societárias, transferências de bens e marcos da investigação	5	15

ID	Descrição	PoC	Homologação
	associados ao Caso, permitindo filtros por tipo de evento, período e nível de relevância.		
7	Marcação de Hipóteses Investigativas: O analista deve poder registrar hipóteses investigativas formais associadas ao Caso, vinculando evidências (registros, documentos, alertas etc.) que as suportam ou contradizem, com acompanhamento do status da hipótese (em análise, confirmada, descartada).	-	15
8	Interface de Chat Analítico por Caso: O sistema deve disponibilizar interface conversacional dedicada por Caso, na qual o usuário formula perguntas em linguagem natural e recebe respostas fundamentadas nos dados do Caso, com indicação das fontes, possibilidade de refinamento da consulta e histórico de perguntas e respostas da sessão.	15	20
9	Sugestões Analíticas Proativas: Com base nos dados já ingeridos no Caso, o sistema deve sugerir proativamente linhas de análise relevantes – como identificação de sócios em comum com investigados de outros Casos (sem expor dados do outro Caso), períodos de movimentação	-	20

ID	Descrição	PoC	Homologação
	financeira atípica ou ausência de declaração fiscal em ano de alta movimentação.		
10	Consulta Cruzada com Acervo Institucional: O usuário deve poder consultar explicitamente a base institucional compartilhada (tipologias de lavagem, manuais técnicos etc.) em linguagem natural, com interface que distingue claramente as respostas provenientes do acervo do Caso das provenientes da base institucional, sem contaminação entre as duas camadas.	10	15
11	Grafo com Narrativa Automática: Ao selecionar um subconjunto de nós no grafo de relacionamentos, o sistema deve gerar automaticamente uma narrativa textual descrevendo os vínculos identificados, os fluxos financeiros entre as entidades selecionadas e os alertas associados, exportável como seção de relatório.	10	20
12	Comparação entre Períodos: Os dashboards de evolução patrimonial e fluxo financeiro devem permitir a comparação visual entre dois períodos definidos pelo usuário, destacando variações absolutas e	5	15

ID	Descrição	PoC	Homologação
	percentuais relevantes e identificando automaticamente os maiores saltos ou quedas no período comparado.		
13	Exportação de Linha do Tempo: O sistema deve permitir exportar a linha do tempo investigativa em formato PDF interativo ou relatório paginado com hiperlinks internos, garantindo a portabilidade e segurança do documento para anexação em petições.	-	10
14	Identificação de Grupos Econômicos de Fato: O sistema deve agrupar automaticamente empresas que, embora possuam CNPJs distintos, compartilhem sócios, endereços, domínios de e-mail ou padrões de movimentação financeira similares, apresentando ao analista o grupo econômico identificado com o grau de similaridade e os critérios utilizados.	10	20
15	Geração de Relatórios: O sistema deve permitir a geração de relatórios técnicos customizados com rastreabilidade das fontes, integridade dos dados e suporte à fundamentação de medidas cautelares, exportáveis em PDF e DOCX.	5	10

ID	Descrição	PoC	Homologação
16	Ingestão e Processamento SIMBA (TXT): Capacidade de processar arquivos TXT no padrão BACEN/SIMBA, realizando o parsing para banco relacional e indexação vetorial dos históricos.	20	25
17	Visualização de Grafo de Relacionamentos: Geração automática de grafo interativo conectando CPFs, CNPJs e transações financeiras a partir dos dados ingeridos.	15	20
18	RAG com Grounding e Citação: As respostas geradas pela IA devem ser baseadas exclusivamente no acervo do Caso, com marcação visual (hiperlink ou nota de rodapé) indicando o documento e a página de origem da informação.	15	20

5. Requisitos Técnicos

ID	Item	Descrição	PoC	Homologação
01	-	Arquitetura Simplificada (MVP): Solução funcional demonstrando fluxos essenciais.	Sim	-
02	6.2.1	Segregação lógica: Isolamento absoluto de dados e interações.	Sim	Sim
03	6.2.2	Segurança da informação: Vedação de credenciais no código, comunicação via HTTPS com TLS 1.2 ou superior, quando aplicável.	-	Sim

ID	Item	Descrição	PoC	Homologação
		Autenticação integrada à Microsoft Entra ID, por meio de protocolos padronizados de autenticação e autorização (ex.: OAuth 2.0 ou equivalentes).		
04	6.2.3	Autenticação Avançada: Configurável por perfil e obrigatória para administradores, quando houver interface administrativa, console de gestão ou operações privilegiadas. Deverá ser compatível com provedores de identidade institucionais.	-	Sim
05	6.2.4	Controle de Acesso (RBAC): Implementação de RBAC para segmentação granular e rastreabilidade.	-	Sim
06	6.2.5	Gestão de Segredos: Cofres (Vaults) com política de rotação de chaves. Aplicável quando a solução realizar armazenamento, processamento ou integração com credenciais, chaves, tokens de acesso ou segredos operacionais sensíveis. Em ambientes on-premise, admite-se o uso de mecanismos institucionais equivalentes do MPRS	-	Sim
07	6.2.6	Proteção de Rede: Defesa contra OWASP Top 10, quando aplicável à arquitetura proposta. Soluções com APIs, acesso externo ou serviços em nuvem deverão	-	Sim

ID	Item	Descrição	PoC	Homologação
		implementar rate limiting e mitigação contra DoS/DDoS. Soluções SaaS, cloud ou híbridas, a proteção de infraestrutura será responsabilidade da proponente.		
08	6.2.7	Arquitetura API First: Disponibilizando interfaces padronizadas e interoperáveis (ex.: REST, GraphQL, gRPC ou equivalentes), com autenticação segura, documentação técnica e versionamento controlado.	Sim	Sim
09	6.2.8	Conteinerização e Serverless: Imagens Docker/OCI com execução portável, permitindo serviços gerenciados/serverless e rollback.	-	Sim
10	6.2.9	Escalabilidade Horizontal: Quando aplicável à arquitetura da solução, balanceamento dinâmico para garantir escalabilidade.	-	Sim
11	6.2.10	Observabilidade: Quando aplicável à arquitetura da solução, emissão de métricas, logs e traços (OpenTelemetry ou equivalente).	-	Sim
12	6.2.11	FinOps (Gestão de Consumo) : Para itens com tarifação variável ou consumo elástico de recursos, limites/quotas parametrizáveis, alertas de consumo, rateio/atribuição e relatórios de auditoria financeira.	-	Sim
13	6.3.1	Trilhas de Auditoria Básicas: Garantia de	-	Sim

ID	Item	Descrição	PoC	Homologação
		integridade e rastreabilidade das operações.		
14	6.3.2	Computação Forense (WORM): Armazenamento de logs em padrão WORM, integridade garantida por hash/criptografia e rastreabilidade da cadeia de custódia.	-	-
15	6.3.3	Inventário SBOM: Documentação SBOM (SPDX ou CycloneDX).	Sim	Sim
16	6.4.1	Prevenção a Injeção e Vazamento: Quando aplicável a soluções que utilizem modelos de Inteligência Artificial generativa ou LLMs, proteção contra manipulação (jailbreak) e vazamento de system prompts.	-	Sim
17	6.4.2	Mitigação de Alucinações (Grounded Generation): Condicionar respostas às fontes e evitar invenção de fatos/normativas.	Sim	Sim
18	6.4.2.1	Metodologia de Aferição Objetiva: durante a avaliação, a solução será submetida a um Dataset de Validação fornecido pelo MPRS.	Sim	Sim
19	6.4.2.2	Critério de Aceite: Quando aplicável a soluções que utilizem modelos de Inteligência Artificial generativa ou LLMs, o desempenho será aferido pela métrica HAM (Hallucination Rate), com HAM máximo de 10% sobre o conjunto de testes.	-	Sim

ID	Item	Descrição	PoC	Homologação
20	6.4.3	Avaliação Contínua e Fine-Tuning: Quando aplicável a soluções baseadas em IA generativa ou RAG, uso de frameworks de qualidade técnica (como RAGAS) medindo fidelidade e consistência.	-	Sim
21	6.4.4	Trilha de Justificativa de IA: Rastreabilidade de ID, fontes/referências, parâmetros da execução e motivos de bloqueio.	Sim	Sim
22	6.4.6	Vedações Operacionais Mínimas: Bloqueio de respostas com "aparência de certeza" sem verificação, vedação de uso de dados reais não autorizados e bloqueio contra desativação de segurança por comandos.	-	Sim
23	6.4.6	Agnosticismo de Modelo de IA: Possibilidade de troca do LLM	-	Sim
24	6.4.7	Gestão de Prompts e Capacitação: Repositório versionado com controle de acesso e fornecimento de manuais/playbooks	-	Sim
25	6.4.8	Agentes Autônomos em Sandbox: Atuação de agentes autônomos em sandbox com limites configuráveis de autonomia e recursos, e bloqueio auditável de operações sensíveis ou escalonamento indevido de privilégios.	-	Sim
26	6.4.9	Métricas de Mídia: Aferição pela métrica WER (Word Error Rate), com WER	-	-

ID	Item	Descrição	PoC	Homologação
		máximo de 10% em conjunto de testes.		
27	6.2.12	Processamento Paralelo e Execução Assíncrona: Quando aplicável à solução, suportar processamento paralelo e execução assíncrona, evitando bloqueios sequenciais e degradação do fluxo principal por chamadas externas ou tarefas de longa duração.	-	Sim
28	6.4.10	Métricas de OCR: Aferição pela métrica CER (Character Error Rate), com CER máximo de 5% em conjunto de testes.	Sim	Sim