Política de Armazenamento de Dados  
Plataforma de Vetorização e Busca Semântica

Versão: 1.0  
Data: 13/09/2025  
Autor: Actionsys  
Empresa: Actionsys Proposal Manager

# Histórico de Revisões

|  |  |  |  |
| --- | --- | --- | --- |
| Data | Versão | Descrição | Autor |
| 13/09/2025 | 1.0 | Versão inicial | Equipe Actionsys |

# 1. Objetivo

Garantir que informações confidenciais de clientes não sejam armazenadas na plataforma, preservando a confidencialidade, integridade e disponibilidade dos dados, em conformidade com normas de segurança da informação (ISO 27001, LGPD).

# 2. Escopo

Aplica-se a todos os dados processados pelo sistema de vetorização e busca semântica, especificamente documentos de clientes (ex.: BRDs, propostas, contratos).

# 3. Princípios

1. Minimização de dados: somente o estritamente necessário será armazenado.  
2. Não retenção de documentos originais: nenhum arquivo (PDF, DOCX, XLSX) será armazenado após a vetorização.  
3. Rastreabilidade segura: será mantido apenas o link/caminho interno para o documento, acessível somente em rede autorizada (VPN).  
4. Proteção criptográfica: embeddings e metadados serão protegidos com controles de banco de dados.  
5. Acesso controlado: apenas perfis autorizados terão acesso à consulta de embeddings e metadados.

# 4. Processo de Armazenamento

1. Upload temporário: O documento original é recebido temporariamente para processamento e, após a vetorização, o arquivo é imediatamente excluído.  
2. Geração de embeddings: Apenas o vetor numérico (embedding) é armazenado no banco; os embeddings não permitem reconstruir o conteúdo original.  
3. Registro de metadados mínimos: nome do documento, caminho interno, ID único, data de processamento.  
4. Exclusão do documento original: Arquivos não permanecem na plataforma, nem em cache.

# 5. Estrutura de Banco de Dados

Tabela: documentos\_embeddings  
- id (chave primária)  
- nome\_arquivo (TEXT)  
- caminho\_rede (TEXT)  
- embedding (vector(1536))  
- data\_insercao (TIMESTAMP)

# 6. Controles de Segurança

- Criptografia em repouso: PostgreSQL com disco criptografado (LUKS ou equivalente).  
- Criptografia em trânsito: conexões obrigatoriamente em TLS/SSL.  
- Controle de acesso: roles no PostgreSQL com privilégios mínimos.  
- Auditoria: logs de quem acessou/consultou embeddings e metadados.  
- Retenção: embeddings e metadados só permanecem enquanto o cliente tiver contrato ativo.

# 7. Exclusão Segura

- Solicitações de exclusão do cliente serão atendidas em até X dias úteis.  
- Embeddings e metadados serão removidos de forma definitiva via DELETE seguido de VACUUM FULL no PostgreSQL.

# 8. Responsabilidades

- Equipe de Desenvolvimento: garantir que o pipeline não armazene documentos originais.  
- Equipe de Operações: manter infraestrutura segura e monitorada.  
- Cliente: prover acesso controlado à rede (VPN) e manter os documentos originais.

# 9. Benefícios para o Cliente

- Garantia de que nenhum documento confidencial sai do ambiente corporativo.  
- Uso de IA apenas sobre embeddings anônimos.  
- Possibilidade de auditoria completa das consultas e acessos.

# Assinaturas

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_  
Responsável Actionsys

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_  
Cliente