

SPRINT 1 — 15/09 A 28/09



**PARDAL**

**USER STORES - DOD E DOR**

André Hideaki Wakugawa

Beatriz A Y Bonatto

Cauê Vieira Da Silva

Gabriel Bartolomeu Guska

Julio Cesar Ferreira De Freitas

Otavio Calderan Bruguel

Paulo Arantes Machado

## Sumário

US01 - Anonimização de Dados Pessoais em Descrições de Chamados.....	2
Descrição .....	2
Lista de Requisitos .....	2
Regras de Negócio para Anonimização .....	2
DOR (Definition of Ready): .....	3
DOD (Definition of Done): .....	3
US02 - Anonimização de Dados Sensíveis (Saúde, Religião etc.) com Análise Semântica (NLP).....	3
Descrição .....	3
Lista de Requisitos .....	3
Regras de Negócio para Anonimização .....	4
DOR (Definition of Ready) .....	4
DOD (Definition of Done) .....	5
US08 - Dashboard de Métricas de Eficiência do Suporte (KPIs) .....	5
Descrição .....	5
Lista de Requisitos (KPIs a Serem Exibidos).....	5
R1 - Total de Tickets Criados .....	5
R2 - Tempo Médio de Resolução (TMR / Average Resolution Time) .....	5
R3 - Taxa de Reincidência (Tickets Reabertos) .....	6
R4 - Conformidade com SLA (SLA Compliance) .....	6
Regras de Negócio e Filtros do Dashboard: .....	6
DOR (Definition of Ready) .....	7
DOD (Definition of Done) .....	7

# US01 - Anonimização de Dados Pessoais em Descrições de Chamados

## Descrição

Eu, como administrador do sistema, desejo que dados pessoais inseridos em campos de título e descrição livre dos chamados sejam automaticamente identificados e substituídos por uma máscara genérica, para evitar o armazenamento indevido de dados pessoais e garantir a conformidade contínua com a LGPD.

## Lista de Requisitos

**R1** - O sistema deverá analisar o conteúdo dos campos de texto Descrição e Título(dos chamados) toda vez que um chamado for criado ou atualizado.

**R2** - O sistema deverá utilizar expressões regulares (Regex) para identificar padrões de dados correspondentes a CPF, endereços de e-mail e números de telefone (formato nacional com DDD).

**R3** - Ao identificar um padrão, o sistema deverá substituir o dado encontrado pela máscara genérica [DADO PESSOAL ANONIMIZADO].

**R4** - O texto original, contendo os dados pessoais, não deverá ser persistido em nenhuma tabela do banco de dados. A anonimização deve ocorrer na camada de serviço, antes da instrução INSERT ou UPDATE.

**R5** - O sistema deverá registrar um log de auditoria na **collection** **\*\*auditLogs\*\*** do **MongoDB** sempre que uma anonimização for realizada em um chamado, informando o ID do chamado e a quantidade de dados que foram mascarados.

## Regras de Negócio para Anonimização

- **Padrões a serem identificados (Regex):**
- **CPF:** (\d{3}[\.\-]? \d{3}[\.\-]? \d{3}[-]? \d{2})
- **E-mail:** ([a-zA-Z0-9.\_-]+@[a-zA-Z0-9.\_-]+\.[a-zA-Z0-9\_-]+)
- **Telefone (Brasil):** (\(? \d{2}\)? \s? \d{4,5}-? \d{4})
- **Máscara de Substituição:** [DADO PESSOAL ANONIMIZADO]
- **Campos Alvo no Banco de Dados (SQL Server):**
- **Tabela:** Chamados, **Coluna:** Descrição (NVARCHAR(MAX))
- **Tabela:** Chamados, **Coluna:** Título (NVARCHAR(MAX))

O dado persistido no SQL Server será automaticamente anonimizado.

## DOR (Definition of Ready):

- **Validação dos Padrões (Regex):** As expressões regulares para identificação dos dados foram validadas e testadas contra diferentes formatos.
- **Mapeamento de Campos Concluído:** Documento que lista todas as tabelas e colunas de texto livre no sistema que podem conter dados pessoais.
- **Definição da Estratégia para Dados Legados:** Decisão de como os dados já existentes serão tratados (confirmando a necessidade da TASK01 de ETL).

## DOD (Definition of Done):

- **Testes Unitários Validados:** Cobertura de testes para o serviço de anonimização.
- **Requisitos Funcionais Implementados:** A lógica de anonimização está integrada nos fluxos de criação e atualização de chamados.
- **Limpeza de Dados Legados:** O script para anonimizar os dados já existentes na base (TASK01) foi executado com sucesso em produção.
- **Pipeline de CI Executado:** O código foi integrado à branch principal e todos os checks passaram com sucesso.

# US02 - Anonimização de Dados Sensíveis (Saúde, Religião etc.) com Análise Semântica (NLP)

## Descrição

Eu, como administrador do sistema, desejo que informações que caracterizem dados pessoais sensíveis (como menções à saúde, religião ou opinião política) inseridas nos chamados sejam automaticamente identificadas via análise de Processamento de Linguagem Natural (NLP) e anonimizadas, para elevar o nível de conformidade com a LGPD e mitigar os riscos associados ao tratamento de dados de alta sensibilidade.

## Lista de Requisitos

**R1** - O sistema deverá analisar o conteúdo das colunas Título, Descrição (tabela Chamados) e Texto (tabela Comentários) toda vez que um chamado for criado ou atualizado.

**R2** - O sistema deverá submeter o conteúdo textual para um serviço externo de Processamento de Linguagem Natural (NLP) para análise semântica.

**R3** - O serviço de NLP deverá ser capaz de classificar trechos do texto em categorias de dados sensíveis, conforme a definição da LGPD (Saúde, Religião, Opinião Política etc.).

**R4** - Se um trecho de texto (ex: uma frase) for classificado com um alto grau de confiança como contendo dado sensível, o sistema deverá substituir a frase inteira pela máscara [DADO SENSÍVEL ANONIMIZADO].

**R5** - O texto original, contendo os dados sensíveis, não deverá ser persistido em nenhuma tabela do banco de dados SQL Server.

**R6** - O sistema deverá registrar um log de auditoria detalhado na collection auditLogs do MongoDB, informando o ID do chamado, a categoria do dado sensível detectado (ex: 'SAUDE') e a pontuação de confiança retornada pela API de NLP.

## Regras de Negócio para Anonimização

- **Categorias Alvo para Detecção:**
  - Dado referente à saúde ou à vida sexual
  - Convicção religiosa
  - Opinião política
  - Filiação a sindicato ou a organização de caráter religioso, filosófico ou político
  - Origem racial ou étnica
- **Serviço de NLP:** A implementação inicial utilizará uma API de mercado (ex: Google Cloud Natural Language API ou Azure Text Analytics for Health) para as funcionalidades de classificação de texto e reconhecimento de entidades.
- **Limiar de Confiança (Confidence Threshold):** A anonimização só será aplicada se a pontuação de confiança da classificação retornada pela API for superior a um limiar configurável no sistema (valor inicial sugerido: **0.85**).
- **Máscara de Substituição:** [DADO SENSÍVEL ANONIMIZADO]
- O dado persistido no SQL Server será automaticamente anonimizado com base na análise contextual do NLP.

## DOR (Definition of Ready)

- **API de NLP Selecionada:** O provedor do serviço de NLP foi escolhido e as chaves de acesso para o ambiente de desenvolvimento estão disponíveis.
- **Limiar de Confiança Definido:** O valor inicial para o confidence score foi definido e aprovado.

## DOD (Definition of Done)

- **Testes Unitários Validados:** Cobertura de testes para o serviço que encapsula a chamada à API de NLP, incluindo mocks da resposta da API para simular diferentes cenários (sucesso, falha, score abaixo do limiar).
- **Auditoria Validada:** Foi verificado que os logs de auditoria estão sendo gravados corretamente no MongoDB, incluindo a categoria do dado e o score de confiança.
- **Pipeline de CI Executado:** O código foi integrado à branch principal e todos os checks passaram com sucesso.

## US08 - Dashboard de Métricas de Eficiência do Suporte (KPIs)

### Descrição

Eu, como gestor, desejo visualizar um dashboard com as principais métricas de desempenho do suporte (KPIs), para monitorar a eficiência da equipe, identificar gargalos, avaliar a conformidade com o SLA e tomar decisões estratégicas baseadas em dados.

### Lista de Requisitos (KPIs a Serem Exibidos)

A seguir, detalhamos cada métrica que deve ser exibida no dashboard.

#### R1 - Total de Tickets Criados

- O que é: Quantidade total de novos tickets abertos em um determinado período.
- Fórmula de Cálculo (Baseada no Schema):  
*COUNT(TicketId) FROM dbo.Tickets*
- Tabela Principal: dbo.Tickets
- Coluna de Contagem: TicketId
- Coluna de Data para Filtro: CreatedAt

#### R2 - Tempo Médio de Resolução (TMR / Average Resolution Time)

- O que é: Tempo médio, em horas ou minutos, que a equipe leva para resolver um ticket desde a sua criação até o fechamento.
- Fórmula de Cálculo (Baseada no Schema):  
*AVG(DATEDIFF(minute, CreatedAt, ClosedAt))*  
para todos os tickets com **ClosedAt IS NOT NULL**.
- Tabela Principal: dbo.Tickets
- Colunas de Cálculo: CreatedAt, ClosedAt

### R3 - Taxa de Reincidência (Tickets Reabertos)

- O que é: Percentual de tickets que foram reabertos após terem sido marcados como "Resolvido" ou "Fechado".
- Fórmula de Cálculo (Baseada no Schema): Identificar o *StatusId* para status de finalização (ex: 'Fechado', 'Resolvido') na tabela *dbo.Statuses*.
- Contar em *dbo.TicketStatusHistory* o número de *tickets* (*COUNT(DISTINCT TicketId)*) que transitaram de um *FromStatusId* finalizado para um *ToStatusId* não finalizado.
- Cálculo: (Total de Tickets Reabertos / Total de Tickets Fechados no período) \* 100.

### R4 - Conformidade com SLA (SLA Compliance)

- O que é: Percentual de tickets que atenderam aos prazos de primeira resposta e de resolução definidos no plano de SLA. Deve ser exibido como dois indicadores separados.
- Fórmula de Cálculo (Baseada no Schema):
  - SLA de Primeira Resposta: Para cada ticket, comparar *DATEDIFF(minute, t.CreatedAt, t.FirstResponseAt)* com o *s.FirstResponseMins* do *dbo.SLA\_Plans* *s* associado ao ticket *t*.
    - Cálculo: (Total de Tickets que atenderam ao SLA de 1ª Resposta / Total de Tickets com 1ª Resposta) \* 100.
  - SLA de Resolução:
    - Para cada ticket, comparar *DATEDIFF(minute, t.CreatedAt, t.ClosedAt)* com o *s.ResolutionMins* do *dbo.SLA\_Plans* *s* associado ao ticket *t*.
      - Cálculo: (Total de Tickets que atenderam ao SLA de Resolução / Total de Tickets Fechados) \* 100.

### Regras de Negócio e Filtros do Dashboard:

**R5** - O dashboard deverá permitir a filtragem de todas as métricas por um **período de tempo** (ex: Hoje, Últimos 7 dias, Mês Atual, Período customizado), com base na data de criação dos tickets (*Tickets.CreatedAt*).

**R6** - O dashboard deverá permitir a filtragem por:

- **Agente:** com base em *Tickets.AssignedAgentId*.
- **Empresa:** com base em *Tickets.CompanyId*.
- **Categoria:** com base em *Tickets.CategoryId*.
- **Prioridade:** com base em *Tickets.PriorityId*.

## DOR (Definition of Ready)

- Stakeholder (gestor) validou as métricas e as fórmulas de cálculo propostas.

## DOD (Definition of Done)

- Um dashboard contendo as 4 métricas principais (R1 a R4) foi desenvolvido.
- Os filtros (R5 e R6) foram implementados e estão funcionando corretamente.
- Os dados exibidos foram validados contra consultas diretas no banco de dados para garantir a precisão.
- O dashboard foi publicado e o acesso foi compartilhado com o gestor.
- O gestor validou o dashboard e confirmou que ele atende às suas necessidades de monitoramento.

